



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 176049

**TO: Ginny Portner
Location: REM/3B02/3C18
Art Unit: 1645
Thursday, January 12, 2006
Case Serial Number: 10/662126**

**From: Toby Port
Location: Biotech-Chem Library
REM-1A59
Phone: 571-272-2523

toby.port@uspto.gov**

Search Notes

Examiner Portner,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

**Toby Port
X22523**



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



STIC-Biotech/ChemLib

176049

From: Portner, Ginny
Sent: Monday, January 09, 2006 4:32 PM
To: STIC-Biotech/ChemLib
Subject: 10/662,126

please search SEQ ID NO 36 and 37; oligomer search both sequences "at least 8 or 10 consecutive amino acids" and "at least 24 to 30 nucleotides". thanks

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02
(571) 272-0862

RECEIVED
JAN - 9 2006
STIC

1/9/2006

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 00:34:19 ; Search time 3013 Seconds
(without alignments)
9055.711 Million cell updates/sec

Title: US-10-662-126-36

Perfect score: 480

Sequence: 1 atggaacaagccatcaaaa.....tgattgcagagagattaa 480

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sta.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_hlg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	480	100.0	480	6	BD009867	BD009867 Proteins,
C 2	480	100.0	13222	1	AE000637	AE000637 Helicobac
C 3	454.4	94.7	14404	1	AE001551	AE001551 Helicobac
4	449.6	93.7	480	1	AY725427	AY725427 Helicobac
5	294	61.3	853	6	AX788995	AX788995 Sequence
6	166	34.6	302212	1	AE017147	AE017147 Helicobac
C 7	165.2	34.4	347961	1	BX571657	BX571657 Wolinella
8	135.2	28.2	314150	1	CJ11168X1	AL139074 Campyloba
9	133.6	27.8	110000	1	CP000025_02	Continuation (3 of
10	113.8	23.7	300029	1	AE015936	AE015936 Clostridi
11	107.6	22.4	34083	1	AF124757	AF124757 Zymomonas
C 12	107.6	22.4	110000	1	AE008692_11	Continuation (12 o
C 13	106.4	22.2	110000	1	AE017180_24	Continuation (25 o
14	103.8	21.6	197615	5	BX470161	BX470161 Zebrafish
15	103.8	21.6	329100	1	SM5591787	AL591787 Snorhizo
16	101.4	21.1	6444	1	AF461795	AF461795 Bartonell
17	101.4	21.1	110000	1	BX897699_07	Continuation (8 of
18	100	20.8	659	6	AX789107	AX789107 Sequence

C 19	99.2	20.7	110000	1	BA000012_04	Continuation (5 of
C 20	99.2	20.7	110000	1	BA000012_05	Continuation (6 of
C 21	98	20.4	110000	1	BX897700_08	Continuation (9 of
C 22	97	20.2	110000	1	BA000040_53	Continuation (54 o
C 23	95.2	19.8	295250	1	BX321862	BX321862 Nitrosomo
24	95	19.8	110000	1	CP000083_16	Continuation (17 o
25	94.8	19.7	110000	1	CP000107_12	Continuation (13 o
C 26	94	19.6	300540	1	AE017154	AE017154 Haemophil
C 27	93.2	19.4	110000	1	AE017321_00	AE017321 Wolbachia
C 28	93	19.4	110000	1	CP000020_21	Continuation (22 o
C 29	92.8	19.3	302040	1	AE017317	AE017317 Desulfochi
C 30	92.6	19.3	10625	1	AE004296	AE004296 Vibrio ch
C 31	92	19.2	348526	1	CR378672	CR378672 Photobact
32	89.6	18.7	110000	1	AE016827_04	Continuation (5 of
33	88.8	18.5	110000	1	CR555306_35	Continuation (36 o
34	88.6	18.5	1755	6	AX414931	AX414931 Sequence
35	88.6	18.5	1927	6	AX416532	AX416532 Sequence
C 36	88.6	18.5	291091	1	AE017330	AE017330 Listeria
C 37	88.6	18.5	324050	1	AL591983	AL591983 Listeria
C 38	88.6	18.5	349980	6	AX641672	AX641672 Sequence
C 39	88.4	18.4	110000	1	AE014291_11	Continuation (12 o
40	88.2	18.4	11371	1	AE006236	AE006236 Pasteurel
41	88	18.3	347071	1	BX640415	BX640415 Bordetell
42	88	18.3	348997	1	BX640427	BX640427 Bordetell
43	88	18.3	349008	1	BX640444	BX640444 Bordetell
C 44	87	18.1	305050	1	AL596173	AL596173 Listeria
45	87	18.1	349980	6	AX417036	AX417036 Sequence

ALIGNMENTS

BD009867 480 bp DNA linear PAT 31-JAN-2002
Proteins, in particular membrane proteins, of Helicobacter pylori,
their preparation and use.

BD009867

ACCESSION BD009867.1 GI:18638240

VERSION JP 2001502886-A/8.

KEYWORDS Helicobacter pylori

SOURCE Helicobacter pylori

ORGANISM Helicobacter pylori

REFERENCE 1 (bases 1 to 480)

Knapp,B., Hundt,E. and Schmidt,K.H.

Proteins, in particular membrane proteins, of Helicobacter pylori,

their preparation and use

Patent: JP 2001502886-A 8 06-MAR-2001;

CHIRON BEHRING GMBH & CO

OS Helicobacter pylori

PN JP 2001502886-A/8

PD 06-MAR-2001

PF 25-JUL-1997 JP 199808651

PI 26-JUL-1996 DE 196 30 390.7

PR BERNHARD KNAPP,ERIKA HUNDT,KARL HEINZ SCHMIDT PC

CJ2N15/31,C07K14/205,C07K16/12,G01N33/53,A61K31/70,A61K39/106, PC

AE1K39/335

CC Strandedness: Single;

CC Topology: Linear;

FT Key Location/Qualifiers

FT source 1..480

FT Location/Qualifiers

1..480

/organism="Helicobacter pylori"

/mol_type="genomic DNA"

/db_xref="taxon:210"

ORIGIN

Query Match 100.0%; Score 480; DB 6; Length 480;

Best Local Similarity 100.0%; Pred. No. 3.8e-116;

Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGAACAAAGCCATCAAACTTGAATCTCAATCTCAATTTTATAGAGCATATCTTACAAAT	60
Db	1	ATGGAACAAAGCCATCAAACTTGAATCTCAATCTCAATTTTATAGAGCATATCTTACAAAT	60
QY	61	CTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTACAGCCATAAA	120
Db	61	CTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTACAGCCATAAA	120
QY	121	AAAAATTGTCGCTTATAAGAAATATCACTTTTAATGAAGACGTGTTTAACGGCGATTTCCCT	180
Db	121	AAAAATTGTCGCTTATAAGAAATATCACTTTTAATGAAGACGTGTTTAACGGCGATTTCCCT	180
QY	181	AATAAGCCCATTTTCCCGGCGCTTTTGATCGTAGAGGCGATGCGCAAAACGGGAGGTTT	240
Db	181	AATAAGCCCATTTTCCCGGCGCTTTTGATCGTAGAGGCGATGCGCAAAACGGGAGGTTT	240
QY	241	TTAGCCTTCACTAGCTTGTGGGGTTTGAACCTCGTAATCGCCAAACAAAAATCGTGTAT	300
Db	241	TTAGCCTTCACTAGCTTGTGGGGTTTGAACCTCGTAATCGCCAAACAAAAATCGTGTAT	300
QY	301	TTTCATGACATTTGATAAGCTTAAATTCGCATCCCTGTAACCCGCGCAGACATTAGAA	360
Db	301	TTTCATGACATTTGATAAGCTTAAATTCGCATCCCTGTAACCCGCGCAGACATTAGAA	360
QY	361	TACCATTTAGAAGTCTTAAAGCATAGGCGCATGATCGGCAAGTGGTGCGCGCTCAA	420
Db	361	TACCATTTAGAAGTCTTAAAGCATAGGCGCATGATCGGCAAGTGGTGCGCGCTCAA	420
QY	421	GTGGATGGCAAGTGTCTGCTGAAGCCGAATGAAAGCCATGATTCAGAGAGAGATTAA	480
Db	421	GTGGATGGCAAGTGTCTGCTGAAGCCGAATGAAAGCCATGATTCAGAGAGAGATTAA	480
RESULT 2			
LOCUS	AE000637/c		
DEFINITION	Helicobacter pylori 26695 section 115 of 134 of the complete genome.		
ACCESSION	AE000637		
VERSION	AE000637.1		
KEYWORDS	GI:2314536		
SOURCE	Helicobacter pylori 26695		
ORGANISM	Helicobacter pylori 26695		
REFERENCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.		
AUTHORS	1 (bases 1 to 13222) Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.		
TITLE	The complete genome sequence of the gastric pathogen Helicobacter pylori		
JOURNAL	Nature 388 (6642), 539-547 (1997)		
PUBLISHED	9252185		
REFERENCE	2 (bases 1 to 13222)		
AUTHORS	Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Cotton, M.D., Weidman, J.M., Fujii, C., Bowman, C., Watthey, L., Wallin, E., Hayes, W.S., Borodovsky, M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-AUG-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
REFERENCE	3 (bases 1 to 13222)		

AUTHORS	White, O.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		
FEATURES	Location/Qualifiers		
source	1. 13222		
	/organism="Helicobacter pylori 26695"		
	/mol_type="genomic DNA"		
	/strain="26695"		
gene	/db_xref="taxon:85962"		
	complement(72..1343)		
CDS	/gene="HP1366"		
	complement(72..1343)		
	/gene="HP1366"		
	/note="similar to SP:P23191 GB:X56977 PID:44182 percent identity: 37.14; identified by sequence similarity; putative"		
	/codon_start=1		
	/transl_table=11		
	/product="type IIS restriction enzyme R protein (MBOIIR)"		
	/protein_id="AAD08410.1"		
	/db_xref="GI:2314537"		
	/translation="MPKLEKILLETOLDPSKECLKFLANRIKSDYRGLHLSORNY DQNKITIIQIFNEVGSGFLQIRTTDMSKRPSNIIEEYAKVVDNICKSEMPQDNS GKNQVTDLSLRKNLFVDMHRMGLIERIKNKEPTNPIYIQSNIKYISLTPLAIEFLNA QDLRNFCYTOALENLQFGAECEVMEIDNHYLDIEEMFVFTFINENFTRSE IIEVGRYSLSRIQEKELKELVODYCNPNHFNKLEKRDYHWNKNGAQOIFSLLEQ SVFEFTNKRLLILKALNEENKNDKKLSRIKELALYFEKHGVKKEGFLHHIVPLC LARSIEPDLDDKWNLIYIDAFNHAKISQTNKHILCYFKNCDDVLSKGLKEQBSL YLTIVENLYKLDLQNALMKIKOLLHLSKNG"		
	complement(1359..2141)		
	/gene="HP1367"		
	complement(1359..2141)		
	/gene="HP1367"		
	/note="similar to SP:P23192 GB:X56977 PID:44181 percent identity: 59.27; identified by sequence similarity; putative"		
	/codon_start=1		
	/transl_table=11		
	/product="type IIS restriction enzyme M1 protein (mod)"		
	/protein_id="AAD08411.1"		
	/db_xref="GI:2314538"		
	/translation="MILNKIYIEDVFTFLDKLEKSDVLAIDIPPYNLIKASWDSFKN DERPLTSPYAWIDMKPLKDTGTFYIPTNPFNCALFLAYLHKVHFLNFTWVKCD GFANAKRYNHAQBSILFYSMHKQNTFNADIEIRIAYESAERIKAQSGIKLNKRW FPNPKGLCLDWELTISQHWKEKGLKPKHSIKPKALIERMIKASSHKNDLILD LFGSGMTSLVAKSLERNFICESHAEYVHGSLEMFYNECE"		
	complement(2128..2991)		
	/gene="HP1368"		
	complement(2128..2991)		
	/gene="HP1368"		
	/note="similar to SP:P29568 percent identity: 32.96; identified by sequence similarity; putative"		
	/codon_start=1		
	/transl_table=11		
	/product="type IIS restriction enzyme M2 protein (mod)"		
	/protein_id="AAD08412.1"		
	/db_xref="GI:2314539"		
	/translation="MNINKPYHSSSTNNNEVPDSDVLIITSPYPYFNKIDYAKNGTQD LQNSAQHVEDIGALEYEDYLLGLLKVWLECYRALKPKGKLCINLYPMLMKVLLNTH LQNSHIFDLHADIQHSILHLDNNMKNPKMFLDDVYIWKRAMPTKRLMFGSPYPRNF YAONTIEFIVGVKDGPKOPTERQEQSOLTQBEWVEFTQIWEIPIPNKNDIAGFK HAALMPAELARLLRLSCVGDVLDPPSGSGTTLREAKLLKRNFIGELYENYKPLI EQKLGUNLDFDE"		
	3240..4664		
	/gene="HP1369"		
	/note="This region contains an authentic frame shift and is not the result of a sequencing artifact; similar to SP:P40814 percent identity: 39.74; identified by sequence similarity; putative; type III restriction enzyme M 4802..5668		
	/gene="HP1370"		

```

CDS
4802..5668
/gene="HP1370"
/translation="similar to GB:L42023 PID:1006312 PID:1221173
PID:1205306 PID:1574609 percent identity: 38.41;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="type III restriction enzyme M protein (mod)"
/protein_id="AAD08413.1"
/db_xref="GI:2314540"
/translation="MYKLDENGFKYKLDPMWAKSGNNFSPYTFNLNGKTSPPSGTFR
YSGTLKDEQNNRIYFNGNPKMAKRYLSEVAGRSKSTPFDGSEVYNGNLNGDAEIKQ
LFNGKVNPKPEALLQRIETSTKENDLVLDFAGSGTTCAVAHKMKRYIGIENG
EPHESVILRLKVIQGFSGAKRPFDDGGVIVKVEYELRLKIKYEDNDDKPLAY
DQYSIDLVKRNESYTLINLEALEKMGVDIKETLENLWGVGFBNFKVVKRFGNDKEV
EILKALKALIIW"
5679..8585
/gene="HP1371"
5679..8585
/gene="HP1371"
/translation="similar to GB:L25415 PID:496157 percent identity:
26.73; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="type III restriction enzyme R protein"
/protein_id="AAD08414.1"
/db_xref="GI:2314541"
/translation="WAKKKDLTTDNEIFVAQKLAELNEINEPLERLDPKSFNN
KELLDYQQOALINAPMLVAYFRDPKKNKFFVAFVQKHYSFAHCDPAKKKLNPLKLS
HKFNHCVSPENFENLAPYMATGSGKTIVILKVELLSVALRGLIPKQIMFPSA
NENLQOFKEIEKYNRNDYFPQIDFKLSKSVTHKDFRPAKPSQVIRKQITLIFYRAD
LWNEDEENKLLNYDNGENYVILDEAHGKNSESRRQAFISLSLKGLFPNPSA
TTTESDLITAVNYISVGEWMLKYGKESVLLKQNNLNAFKELKOLNREKIALIKA
LILLGQKRYTGVYFPLMLVTHSVNVKNSDAEIPFKTLARVLENDGDFLAK
EDLLELKNPELPSDDKDKVVKVFEKGLKSMOPKGLKEVFYANNCHIIVLINPK
NLQVAFPLNTSDKVFCLIRIGDITWICEKLSKVSVKSNLSPKEESYFSQIDKSSI
NIVGSRPTDQTSRPSVILFNLIGLDDAKLVKQSGRGVRLIESVKNQRQLAY
LDIDGAILKALPANAAMLETFVIPNTYASLAILFKQESKESNGRSWREIKLEK
TPIKHALFVPCYKQTSILEIPESAFKMSKFNKDLKKEFNLMSEKHFILKHEIYD
PKYMQLKQTOEAPFNKVSWHYKQLDYMSIISIKGLVPNOKVPKDBFNALDSKIV
HKPRILKVADKKEVLPKTOEYKVAPLDKETLIIKLAGOEIDPDYDTEKHONKTFK
GSAELLKLEHYTPIIKAKNDWLKHVVVSESDPFBELLKIETLEQNTDFWAPS
KIDEHLNDLPIFYNNAAERKFFDFIFWLEKGGTQIICFIDPKGSKHFDYBHKADAY
QIFDKIKFNPKDNLKIKVILKFKYGDKEVADGYDYIWKGLKLEDFFLKOLA"
complement (8635..9381)
/gene="HP1372"
complement (8635..9381)
/gene="HP1372"
/translation="rod shape-determining protein (mreC)"
/product="rod shape-determining protein (mreC)"
/protein_id="AAD08415.1"
/db_xref="GI:2314542"
/translation="WRFYKFLMLLGIPIFYFLDPKGSYSVIDRIKNAALNKNLS
LDNVQAYFQAOIKFQERLILKALKNADKLERLNSIYPLENPMYTFPTWTS
FISLDETSVNLNVEENIKYGLVSHNQAIGIAVLEKGRGLNGFLNAHKRCAYSVM
IGQNVGLGIFGNFQKQELVVDPIVPSAEINIGDQVLTSGLDIGFAGVFGVSVSIED
HYTKSAVLKNAFLSKALRVFLSDVKN"
complement (9385..10428)
/gene="HP1373"
complement (9385..10428)
/gene="HP1373"
/translation="similar to SP:P13519 GB:M22055 PID:606191
PID:1128967 GB:U00096 percent identity: 51.87; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="rod shape-determining protein (mreB)"
/protein_id="AAD08416.1"
/db_xref="GI:2314543"

gene
CDS
/translation="WIRSKLIGLFSHDIAIDLTANTIVLVKGGQIIINEPSIVAVRM
GLFSPKAYDILAVGSEAKEMLGKTPNSRAIRPMKQGVIAVDYITAKMIRPIEKVHK
RKTWIRPIWCVPIGLTSVERNAVKESTLSAGAREVFLIEPEPMAAIGAGLPVKEPQ
GSLIVDIGGTEIGVISLGVISKRSIVAGDKLDQSVIVRYIRKFNLLIGERTGE
IKTEIGCAIKPLPLTMEVSRDQVSGLLHTIELSSDDVFEAIKQDQVREISALRSVL
BEVKPDLAKDIOVQGVVLITGGGALIKGLDKVLSLDMVLPVYVVDPEPLAVAKGTGEAI
QDLDLLSRVGFSE"
complement (10471..11811)
/gene="HP1374"
complement (10471..11811)
/gene="HP1374"
/translation="similar to GB:L18867 SP:P3138 GB:Z23278 PID:347729
PID:388977 percent identity: 55.04; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="ATP-dependent protease ATPase subunit (c1px)"

Query Match 100.0%; Score 480; DB 1; Length 13222;
Best Local Similarity 100.0%; Pred. No. 3.3e-116;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAACAAAGCCATCAAACTTCGAATCTCAATTTTATAGAGCATATCTTACAAATT 60
DB 13108 ATGGAACAAAGCCATCAAACTTCGAATCTCAATTTTATAGAGCATATCTTACAAATT 13049
QY 61 CTACCTCACCGCTATCCATGCTTTTATAGTATAGAAATTTATAGAGTTTACAAGCCATAA 120
DB 13048 CTACCTCACCGCTATCCATGCTTTTATAGTATAGAAATTTATAGAGTTTACAAGCCATAA 12989
QY 121 AAAATTGTCGCTTATAGAAATATACATTTTATGAAGACGTTTAAACGGGCAATTTCCCT 180
DB 12988 AAAATTGTCGCTTATAGAAATATACATTTTATGAAGACGTTTAAACGGGCAATTTCCCT 12929
QY 181 AATAAGCCCAATTTCCCGGGGTTTTCATCTGATAGAGGCGATGCGCAACCGGAGGTTT 240
DB 12928 AATAAGCCCAATTTCCCGGGGTTTTCATCTGATAGAGGCGATGCGCAACCGGAGGTTT 12869
QY 241 TTAGCCCTTCACTAGCTGTGGGGTTTGACCTCAAAATCGCAAAACAAAAATCTGTAT 300
DB 12868 TTAGCCCTTCACTAGCTGTGGGGTTTGACCTCAAAATCGCAAAACAAAAATCTGTAT 12809
QY 301 TTGATGACGATTGATAGGTTAAATTCGATCCCTCTGTAACCCAGGCGACAGATTAGAA 360
DB 12808 TTGATGACGATTGATAGGTTAAATTCGATCCCTCTGTAACCCAGGCGACAGATTAGAA 12749
QY 361 TACCATTTAGAACTTTAAAGCATAGGGCATGATCTGGCAAGTGGTGGCACGGCTCAA 420
DB 12748 TACCATTTAGAACTTTAAAGCATAGGGCATGATCTGGCAAGTGGTGGCACGGCTCAA 12689
QY 421 GTGGATGGCAAGTGGTGGTGGTGAAGCCGAATTGAAAGCCATGATTGAGAGAGATTAA 480
DB 12688 GTGGATGGCAAGTGGTGGTGGTGAAGCCGAATTGAAAGCCATGATTGAGAGAGATTAA 12629

RESULT 3
AE001551/c
LOCUS AE001551
DEFINITION Helicobacter pylori, strain J99 section 112 of 132 of the complete genome.
ACCESSION AE001551
VERSION AE001551.1
KEYWORDS GI:4155887
SOURCE Helicobacter pylori J99
ORGANISM Helicobacter pylori J99
REFERENCE 1 (bases 1 to 14404)
AUTHORS Alm,R.A., Ling,L.S., Moir,D.T., King,B.L., Brown,E.D., Doig,P.C., Smith,D.R., Noonan,B., Guild,B.C., deJonge,B.D., Carmel,G., Tummino,P.J., Caruso,A., Uria-Nickelsen,M., Mills,D.M., Ives,C., Gibson,R., Merberg,D., Mills,S.D., Jiang,Q., Taylor,D.E., Vovis,G.F. and Trust,T.J.

```

TITLE Genomic-sequence comparison of two unrelated isolates of the human
JOURNAL Gastric pathogen Helicobacter pylori
PUBMED Nature 397 (6715), 176-180 (1999)
REFERENCE 9923682
AUTHORS 2 (bases 1 to 14404)
TITLE King,B.L., Alm,R.A. and Trust,T.J.
JOURNAL Direct Submission
COMMENT Submitted (12-JAN-1999) Astra Research Center Boston, 128 Sidney
Street, Cambridge, MA 02139, USA
Address all correspondence to: hpc@arcb.us.astra.com or Richard
A. Alm, Astra Research Center Boston, 128 Sidney Street, Cambridge,
MA, 02139. Lo-See L. Ling, Donald T. Moir, Douglas R. Smith,
Braydon C. Guild, Gilles Carmel, Anthony Caruso, Debra M. Mills,
Rene Gibson, and Gerald F. Vovis are with Genome Therapeutics
Corporation, 100 Beaver Street, Waltham, MA, 02453. Qin Jiang and
Diane E. Taylor are with the University of Alberta Department of
Medical Microbiology and Immunology, Edmonton, Alberta, Canada, T6G
2H7 and the Canadian Bacterial Diseases Network. All other authors
are with Astra Research Center Boston, 128 Sidney Street,
Cambridge, MA, 02139. Putative identifications, sequence
alignments, and name and sequence search capability are available
at ARCB's World wide Web site. (URL:
http://www.astra-boston.com/hpylori).

FEATURES
source

1. .14404
/organism="Helicobacter pylori J99"
/mol_type="genomic DNA"
/strain="J99"
/db_xref="taxon:85963"
complement (71. .1531)
/gene="dnaB"
/note="synonym: jhp1280"
complement (71. .1531)
/gene="dnaB"
/note="similar to H. pylori 26695 gene HP1362"
/codon_start=1
/transl_table=11
/product="REPLICATIVE DNA HELICASE"
/protein_id="AAD06865.1"
/db_xref="GI:4155895"

/translation="MDHLKHLQOLQNIERIVLSGIVLANHKEIHSVLEPSDFYPP
HGLFEALKLHEVNCIPIDENIRQMPKDKQISEDLLVAIPAASPIDNIEAYEISIK
NASIKRRLTANTREQALESAQSDILNAVEREYVALLNGSTIEGFRGKEVLES
TNWLTENQKSLTKVTGPTGTVQVLDNTSGNQSLVILGARPMSGKTSIMNMVL
SALNDRGVAVSLEMSAQLARLSDLTSLNMHDELSARDDDQWENLAKCFDHL
QKLLPFYKSVYRMDQIRLKLKSKQHKELGAFIDYQLQMSGNATKERHEQIAEI
SHEKTLARELEIPIIALVOLNRSLRDRDKPILSDIKDSGGIEQADIVLFLYRGY
IYQMRADNKIDIKKEGVBEAQELHLKVNREIRHKQNGSIEEAIIIVAKNRNGAT
GTVYTRFNAPFTRYEDMPVDSHLEEQETKFEMPTT"

complement (1542. .2936)
/gene="jhp1281"
complement (1542. .2936)

/gene="jhp1281"
/note="similar to H. pylori 26695 gene HP1363"
/codon_start=1

/transl_table=11
/product="putative"
/protein_id="AAD06866.1"
/db_xref="GI:4155896"

/translation="MLSVTEKGNALDRVLEWLLSEDILMENAAMALERAVLQNASL
GARVITLCSGONGGPGYTLARLVGRFKTLVPEMKLAKSPMCQLOKERAKVGVVVK
ANEEKNEDECVDVDCVVGSAFKGLEPDPESISOKARFKIACDIPSGIDKGRV
DKRAFADTTISGAILKSLSDKADYIGELKVLGVFNQIYEIPTDPLLEKSDL
KIPLDRKNAHGDYGAHVLLGSGAGLLSALSFGSVVTSIQALECEITSNKK
VLETPCEFPKSAFGLMGLENTPKDFKWELEAPCVLDGAVFYHKEVLAQKE
PIELTHPEKFLSLKSGVGINISMLELDNKLKLEIARDPSQYPRKVVLLKANTLIAHQ
GRVFINLGSVALAKAGSDVLGVLVLSLQNTYPLDAAINASLAHALAGLEFKNHY
ALTPDLLEIKREL"

complement (2939. .3973)
/gene="jhp1282"
complement (2939. .3973)

/gene="jhp1282"
/note="similar to H. pylori 26695 gene HP1364"

/codon_start=1
/transl_table=11
/product="putative histidine kinase sensor protein"
/protein_id="AAD06867.1"
/db_xref="GI:4155897"
/translation="MASKIASEIILQHMOTHADYHNALNALISRYKDASTALFDSKKE
VLYSNIPESADLIKTHKEGFFSGEYIYFSDTEFAHLGVAKMLPKNKRPLHPSLY
RNMVLYVFAFLCVIGSVFLGRFLKRNRIIRIDHFLGVAKMLPKNKRPLHPSLY
KTELEQOHRRIKIAIQRMFLYSLSYVMQDIERESFALLDLKALITKENTLFSEM
IDYHKLKSDLVGVVAKAKQDFISLYSNLMAAIKYSVMHGVIHIELTPEFLKVKN
LGVEIPDKITKLSVRYARNSSVLGVGIGLDLVKKVCKVKMRLEIHSBPSLKSFPH
ENSFCVQFQG"
complement (4107. .4748)
/note="similar to H. pylori 26695 gene HP1365"
/codon_start=1
/transl_table=11
/product="putative TRANSCRIPTIONAL REGULATOR"
/protein_id="AAD06868.1"
/db_xref="GI:4165002"
/translation="MQKIFLDEDDYLLSESIFEHLGYEYVCFAPNGKEAHERLSV
ERNFLLDVQVPEMNSLELFRKIKNDFLISTPVIPTALQDNATLKNAFLNGASDYL
KKPFDLELEARKIRFNDPDIEMPNIFYQHALNVKGEKILAPKTAQLLEYFLEH
KQQLISSQALENNLWEQAIDDSLTRTVIKVLRKLKGNKCIETHKGVGVRFNPL"
complement (4110. .4748)
/gene="jhp1283"
4857. .6806
/gene="jhp1284"
4857. .6806
/gene="jhp1284"
/codon_start=1
/transl_table=11
/product="TYPE II DNA MODIFICATION ENZYME
(METHYLTRANSFERASE)"
/protein_id="AAD06858.1"
/db_xref="GI:4155888"
/translation="MKSETTYKDFCELSGILFNFDTSLLLESKYNKEKVDILFDLKS
VQDKTKNTLNFSTLSSKGTQTKNLEKCSNQGVKLDDEILKKAFAVFKKQGSW
ELYIKHNAQFLKEQDLVLYFLFKEMTAFDHKRLNGINTIKVLEALEVALVSEFEN
ELCIHNKPRILVNLHPIVSLDKAKNYDLSKITNHPNYPKVQKVEDLNLEADNL
LGNEFLPDLTYFKDLEEEVKSIPNEINGTLIKSENYOALNSLKRYKEAIDCIYI
DPPTGSDPAYIDRFQDSTWSLMNRNLQAVDFLSPOGNFYLLHLYRANLYGEMLL
NDIFSKENFNEIWHFRTYQGOIQSNFPRKSHLSLWISKNCNVNPFRTYSDNYK
TVDYRREFIVDNKNIVPNYPKADSRFDYLRLOSTKEPKNGDIIATINGVID
DWTVDIAIDKADLERLOGLTKQPEKLERIISKASNSNSIVCDFPAGSGTTCAVD
HKLKRYIGVGEHEPESVILPRLLKVIQVKGSGALKKEFNGGGVIVKVVLESYBEILR
KIKYENDKPLAVEEQVSLVERKNESYTLNIEALENGVDIKETLENLHGVGVFEFFN
EKVKFGKNDKESVILKALKEALW"
6816. .9734
/gene="jhp1285"
6816. .9734
/gene="jhp1285"
/note="similar to H. pylori 26695 gene HP1371"
/codon_start=1
/transl_table=11
/product="putative"
/protein_id="AAD06859.1"
/db_xref="GI:4155889"
/translation="MAKKQEVNRNEIFVAQKLAEEELNANEINDPLEMLDPKSPDDN
KELDYQOQALINAFRVLYAFRDFKESKKEFYAQKHSFANCDFAKKLNLHLKS
HFYKHCVFENFENIRLAFYMATGSKTIVIKVELLSVAMGMLIPKKNIMFSA
NEHLIKQFEKIEKRNKDYQSDIDFNKLSVKNKDYFAPKDSLMKEIALFYFRAD
LMSDEESKNLLNYDCWDNGENVVILDEAHKGNKTESKQAIIFSLLSLKGFLNPSA
TYTEESDLITAVNLSYGVKESVLLKNNLNAFKELDNLNREKHEIALKA
LLLLQMKRYKRGYFHDPLMLPFTHSVNMNSDAEIFFKTLARVIENDSDSPKAK
DOLLBEUKNEFLPSDGOKKQKQKIVFKEKSGMDPKGLKEAVFYASNGHIEVIN
PKNNKTEAFKNSDKVFCILIRIGDITETWEKLSKVSVKSNISFKEBSYDIKNS
SINILVGSRAFDGWDSTPVSILFLNLGDDDAKLKLVKDFGRGVIESVKNQRRL
AYLEIDRAIKDKLPNAAEMLEFVPTNHASLEALKFOKESNGENRGPRRIKL
EKTRIEHALFVPCYKEQTNALKISQASFKMSKFNKDLKEHFLMSEKHEILAKHEI
YDPKRYALLKEMIQTAFKKVSTWHYKOLDYMISEIKGKLYPNQKVPKDFNADNEK
IVHFRKVKADKKEKLIQIOEVKEHADPLDKETLRKKIAGIEQIDPYDEKHKQRTF
KVGDAELLKKEHYTPLIKAKNCDWLKVVVKSESDFBELKLTETLQENYDFWA
FSKIDEHLNLFITYIDNATERPPDFIFWLQKGGTQIIICFIQPKGSKHTDYEHKAD

gene

gene

CDS

gene

CDS

gene

CDS

gene	AYQLPEDKVFNPCKDDPNLKI.KVVLKPYGDKDDVGBRYRDLWIEKGLKLEYFFPSGGFKGLA	
CDS	complement (9784. .10530)	
	/gene="jhp1286"	
	complement (9784. .10530)	
	/gene="jhp1286"	
	/note="similar to H. pylori 26695 gene HP1372"	
	/codon_start=1	
gene	/transl_table=11	
	/product="putative"	
	/protein_id="AAD06860.1"	
	/db_xref="GI:4155890"	
	/translation="MRFPFKFWLLIGILFIYFLDIKSSSYISDRVKSLAMSAKNSL	
	LDNVQAYFQAQNIKEFKERLILERALKLENADLKERLNSIYPLENPKTYTPTFMIS	
CDS	FINLEDTHSVNLNIVNLEENKIYGLVSHNQDAIGIAVLEKRLNGFLNNAHKRCAYSVM	
	IQGNQVLGFTGNFKQELVPIFVPSAEINIQDVLTSGLDIFGAGVFGVESSVED	
	HYTKSAVLKNAPLSEAKLRHVFLSGVKN"	
	complement (10534. .11577)	
	/gene="mreb"	
	/note="synonym: jhp1287"	
gene	complement (10534. .11577)	
	/gene="mreb"	
	/note="similar to H. pylori 26695 gene HP1373"	
	/codon_start=1	
	/transl_table=11	
	/product="ROD SHAPE-DETERMINING PROTEIN"	
CDS	/protein_id="AAD06861.1"	
	/db_xref="GI:4155891"	
	/translation="MFPSKLIGLFSHDIAIDLTANTIVLVKGQGIINEPSIVAVRM	
	GLFDSKAVDILAVGSEAKEMLGKTPNSIPAIRPKDGVIAVDITAKMIRYPIEKAKH	
	RKTWRPRIMVCVPYGLTSVERNAVESALSAGAREVLIEPMAAIGAGUPVKEPQ	
	GSLIVDIGGTEIGVISLGGIVISKSIRVAGDKLDQSIYIRKPNLLIGERTGEE	
gene	IKETICAIKLDPPLTMEVSGRDQVSLHTIELSDSDVFEAIKQDREISSALRSVL	
	BEVKPDLVRDVIQONGVWLTGGGALIKGLDKYLSMDMVKLPVTVGDPELLAVAKGTGEAI	
	QDLDLLSRVGPE"	
	complement (11630. .12988)	
	/gene="clpx"	
	/note="synonym: jhp1288"	
Query Match	94.7%; Score 454.4; DB 1; Length 14404;	
Best Local Similarity	96.7%; Pred. No. 2e-109;	
Matches 464; Conservative	0; Mismatches 16; Indels 0; Gaps 0;	
QY	1	ATGGAACAAGCCATCAAAACTTGCATCTCAATCTCAATTTTTATAGAGCATATCTTACAAATT 60
Db	14284	ATGGAACAAGCCATCAAAACTTCAATCTCAATTTTTATAGAGCATATCTTACAGATT 14225
QY	61	CTACTCACCCTATCCCATGCTTTTGTGATAGATAGAAATATAGAGTTACAAGCCAAATAA 120
Db	14224	TTACTCACCCTATCCCATGCTCTTAGTGATAGGATATATAGAGTTACAAGCCAAATCAA 14165
QY	121	AAAATTTGTCGTTTATAAGAAATATCACTTTTAAATGAAGACGTTTAAACGGGCAATTCCT 180
Db	14164	AAAATTTGTCGTTTATAAGAAATATCACTTTTAAATGAAGATGTTTAAACGGGCAATTCCT 14105
QY	181	AATAAGCCCATTTTCCCGGGCGTTTGTATCGTAGAGGCGATGGCGCAACACGGAGGGTTT 240
Db	14104	AATAAGCCCATTTTCCCGGGCGTTTGTATCGTAGAGGCGATGGCGCAACACGGAGGGTTT 14045
QY	241	TTAGCCTTCACTAGCTTGTGGGGGTTTGAACCTGAAATCGCCAAACCAAAATCGTGAT 300
Db	14044	TTAGCTTTCACTAGCTTGTGGGGGTTTGAACCTGAAATAGCCAAACCAAAATCGTGAT 13985
QY	301	TTATGAGCATTTGATAGAGTTAAATTCGCGATCCCTGTAAACCCAGGCGACAGATTAGAA 360
Db	13984	TTATGAGCATTTGATAGAGTTAAATTCGCGATTCCTGTAAACCCAGGCGACAGATTAGAA 13925
QY	361	TACCAATTTAGAGTCTTAAAGCATAGGCGCATGATCTGGCAAGTGGGCGGCGCTCAA 420
Db	13924	TACCAATTTAGAGTCTTAAAGCATAGGCGCATGATCTGGCAAGTGGGCGGCGCTCAA 13865
QY	421	GTGGATGSCAAAGTGGTGGCTGGAAGCCGAAATGGAAGCCATGATTTGACAGAGAGATTAA 480

Db	13864	GTGATGCAAGTGGTGGCTGTAAGCCGAATTAAAGCCATGATCGCAGAGAGATAA	13805
RESULT 4			
AY725427			
LOCUS	AY725427	480 bp	DNA linear BCT 08-JUL-2005
DEFINITION	Helicobacter pylori (3R)-hydroxymyristoyl-acyl carrier protein dehydratase (fabZ) gene, complete cds.		
ACCESSION	AY725427		
VERSION	AY725427.1		
KEYWORDS	GI:56684724		
SOURCE	Helicobacter pylori		
ORGANISM	Helicobacter pylori		
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.			
REFERENCE	1 (bases 1 to 480)		
AUTHORS	Liu, W., Luo, C., Han, C., Peng, S., Yang, Y., Yue, J., Shen, X. and Jiang, H.		
TITLE	A new beta-hydroxyacyl-acyl carrier protein dehydratase (FabZ) from Helicobacter pylori: Molecular cloning, enzymatic characterization, and structural modeling		
JOURNAL	Biochem. Biophys. Res. Commun. 333 (4), 1078-1086 (2005)		
PUBMED	15967411		
REFERENCE	2 (bases 1 to 480)		
AUTHORS	Liu, W., Han, C., Wang, Q., Jiang, H. and Shen, X.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-AUG-2004) Drug Discovery and Design Center, Shanghai Institute of Materia Medica, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Zuchongzhi Road 555, Shanghai 201203, China		
FEATURES			
source			
1. .480			
/organism="Helicobacter pylori"			
/mol_type="genomic DNA"			
/strain="SS1"			
/db_xref="taxon:210"			
gene	1. .480		
/gene="fabZ"			
CDS	1. .480		
/gene="fabZ"			
/function="fatty acid and phosphatidic acid biosynthesis"			
/note="enzyme"			
/codon_start=1			
/transl_table=11			
/product="(3R)-hydroxymyristoyl-acyl carrier protein dehydratase"			
/protein_id="AAW22049.1"			
/db_xref="GI:56684725"			
/translation="MEQSHNLQSQFFIEHILQILPHRYPMLLVDRITELQANQIVFA YKNITFNEVDVFNHFPNKPFPGLVIEGMAQSGFLAFTSLWGFDFEIAKTIKIVFM TIDKVKFRIPVTPGDRLEHLEVLKHGMWQVGGTAQVDGKVAEELKAMIAERE"			
ORIGIN			
Query Match 93.7%; Score 449.6; DB 1; Length 480;			
Best Local Similarity 96.0%; Pred. No. 4.3e-108;			
Matches 461; Conservative 0; Mismatches 19; Indels 0; Gaps 0;			
Qy	1	ATGGAACAAGCCATCAAACTTGCATCTCAATTTTTATAGAGCATATCTTACAAATT	60
Db	1	ATGGAACAAGCCATCAAACTTGCATCTCAATTTTTATAGAGCATATCTTACAGATT	60
Qy	61	CTACTCACCCTATCCCATGCTTTTGTGATAGATATAGAGTTACAAGCCCAATAAA	120
Db	61	TTACTCACCCTATCCCATGCTCTTAGTGATAGATATACAGATTACAAGCCCAATCAA	120
Qy	121	AAAATTTGCGTTATAGAAATATCACTTTTAAATGAAGACCGTGTTTTAAACGGGCAATTC	180
Db	121	AAAATTTGCGTTATAGAAATATCACTTTTAAATGAAGACCGTGTTTTAAATGGCAATTC	180
Qy	181	AATAAGCCCATTTTCCCGGCGTTTGTATCGTAGAGGCGATGGCGCAACCGGAGGGTTT	240
Db	181	AATAAGCCCATTTTCCCGGCGTTTGTATCGTAGAGGCGATGGCGCAACCGGAGGGTTT	240

```
QY 241 TTAGCTTCACCTAGCTGTGGGGGTTTGACCTGAAATCGCAATCGCCAAACAAAAATCGTGTAT 300
Db 241 TTGGCTTTCACCTAGCTGTGGGGGTTTGACCTGAAATAGCCAAACAAAAATCGTGTAT 300
QY 301 TTCATGACGATTTGATAAGGTTAAATTCGCATCCCTGTAACCCCGACGCGCACAGATTAGAA 360
Db 301 TTCATGACGATTTGATAAGGTTAAATTCGCATCCCTGTAACCCCGACGCGCACAGATTAGAA 360
QY 361 TACCAATTTAGAAAGTCTTAAAGCATAGGCGCATGATCGGCAAGTGGGTGGCGCGCTCAA 420
Db 361 TACCAATTTAGAGTCTTAAACATAAGGCGCATGATCGGCAAGTGGGTGGCGCGCTCAA 420
QY 421 GTGdATGGCAAGTGTCTGCTGAAGCCGAATTTGAAGCCATGATTCGACAGAGAGATTAA 480
Db 421 GTGGATGGCAAGTGTCTGCTGAAGCCGAATTTGAAGCCATGATTCGACAGAGAGATTAA 480

RESULT 5
AX788995 LOCUS AX788995 853 bp DNA linear PAT 17-JUL-2003
DEFINITION Sequence 1459 from Patent WO02066501.
ACCESSION AX788995
VERSION AX788995.1 GI:32955341
KEYWORDS Helicobacter pylori
SOURCE Helicobacter pylori
ORGANISM Helicobacter pylori
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1
AUTHORS Legrain,P., Rain,J.C., Colland,F., de Reuse,H. and Labigne,A.
TITLE Protein-protein interactions in Helicobacter pylori
JOURNAL Patent: WO 02066501-A 1459 29-AUG-2002;
Hybrigenics (PR) : INSTITUT PASTEUR (FR)
FEATURES
    source      : 1. .853
                /organism="Helicobacter pylori"
                /mol_type="unassigned DNA"
                /db_xref="taxon:210"
ORIGIN
Query Match 61.3%; Score 294; DB 6; Length 853;
Best Local Similarity 100.0%; Pred. No. 7e-67;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCGAACAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
Db 560 ATGGAACAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 619
QY 61 CTACCTCACCGCTATCCCATGCTTTTAGTGATAGAAATTATAGAGTTACAAGCCATAAA 120
Db 620 CTACCTCACCGCTATCCCATGCTTTTAGTGATAGAAATTATAGAGTTACAAGCCATAAA 679
QY 121 AAAATTGCTTATAGAAATATCACTTTAATGAAGAGCGTGTAAACGGGCAATTCCT 180
Db 680 AAAATTGCTTATAGAAATATCACTTTAATGAAGAGCGTGTAAACGGGCAATTCCT 739
QY 181 AATAAGCCCAATTTCCGGCGGTTTTCATCGTAGAGGCGATCGCAACGGGAGGTTT 240
Db 740 AATAAGCCCAATTTCCGGCGGTTTTCATCGTAGAGGCGATCGCAACGGGAGGTTT 799
QY 241 TTAGCCTTCACTAGCTGTGGGGGTTTGACCTGAAATCGCCAAACAAAAATC 294
Db 800 TTAGCCTTCACTAGCTGTGGGGGTTTGACCTGAAATCGCCAAACAAAAATC 853

RESULT 6
AE017147 LOCUS AE017147 302212 bp DNA linear BCT 26-JUN-2003
DEFINITION Helicobacter hepaticus ATCC 51449 section 4 of 6 of the complete
genome.
ACCESSION AE017147 AE017125
VERSION AE017147.1 GI:32262489
KEYWORDS
```

```
SOURCE ORGANISM
Helicobacter hepaticus ATCC 51449
Helicobacter hepaticus ATCC 51449
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 302212)
AUTHORS Suerbaum,S., Jogenhans,C., Sterzenbach,T., Drescher,B., Brandt,P.,
Bell,M., Droge,M., Partmann,B., Fischer,H.P., Ge,Z., Horster,A.,
Holland,R., Klein,K., Konig,J., Macko,L., Mendz,G.L., Nyakatura,G.,
Schauer,D.B., Shen,Z., Weber,J., Frosch,M. and Fox,J.G.
TITLE The complete genome sequence of the carcinogenic bacterium
Helicobacter hepaticus
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7901-7906 (2003)
PUBMED 12810954
REFERENCE 2 (bases 1 to 302212)
AUTHORS Drescher,B. and Suerbaum,S.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2003) Institute of Hygiene and Microbiology,
University of Wuerzburg, Josef Schneider Str. 2, Wuerzburg D-97080,
Germany
FEATURES
    source      : 1. 302212
                /organism="Helicobacter hepaticus ATCC 51449"
                /mol_type="genomic DNA"
                /strain="ATCC 51449"
                /db_xref="ATCC:51449"
                /db_xref="taxon:235279"
                complement (68. .1411)
                /genes="ffh"
                /locus_tag="HH0940"
                complement (68. .1411)
                /genes="ffh"
                /locus_tag="HH0940"
                /codon_start=1
                /transl_table=11
                /product="signal recognition particle protein"
                /protein_id="AAP77537.1"
                /db_xref="GI:32262490"
                /translation="MFNTLTSPFKGIINKIRFSDDEKSLQACDELKTKLKNQVYHK
                AVKQIIQEVQORTKEAGIKGFONFINALQDALSKILQSKSYGSFSPPTPSVIIMMGL
                QSGSKTSSAKALYFKQGGKVLVACDLRLAAI EQLTQAQS IVDI FTPSHSQ
                PNNVYVAKAKQKATEAHYDIVIIDSAGRLAIDEVLMSELEIKRAVASECLYVAD
                SLGQGIQRAQHPDDKIGIDVILSKFSDTKGGIALSIAYIGICIPKFIGSGEKIA
                DFDFIPDRIISRLMGAGDIASLAKTANVINBEAKTIARKIKKGQFGFDFIAQIE
                NVFKLSSMSIASMIPLGGMNALKDVIDLKSGETIKIRAMVNSMTKKRENPDLLN
                GSRKRIAQSGGLEVSIDINRIIKQFQAAKFAKKLSQKGGMQELMSMGNLRIPK"
                1447. 1557
                /locus_tag="HH0941"
                1447. 1557
                /locus_tag="HH0941"
                /note="predicted membrane-associated"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="AAP77538.1"
                /db_xref="GI:32262491"
                /translation="MFNRKFYTKITFTMLLPASNIYDYKILYIFPKDYTK"
                1554. 1964
                /locus_tag="HH0942"
                1554. 1964
                /locus_tag="HH0942"
                /note="FurR family"
                /codon_start=1
                /transl_table=11
                /product="transcriptional regulator"
                /protein_id="AAP77539.1"
                /db_xref="GI:32262492"
                /translation="MMNFEQHLRDKHLKITPQRIATLNQIYNNGHMSIEEYEQIKQI
                YPSISLATIYKVNVALCKANTLREIKAPKDKQKVELSSDKHLHVYCEKGRLLDDIKLD
                TRALEQNCSSAGSYTISDISAVLMGVCPDCKNAS"
                complement (1969. .4014)
                /locus_tag="HH0943"
                complement (1969. .4014)
                /locus_tag="HH0943"
```

```

/notes="similar to Helicobacter pylori JHP0417"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP77540.1"
/db_xref="GI:32262493"
/transl_table=11
/notes="MDNAROTQPTSLRAICQSGALHLLSYAKRYALNNMFPASII
NPKLPEELKSPQAVNYVNCNGEDLANIIDVHTKNFVYPTKKGKGLKHTMLEONFALSS
NPLNSAYTLHNRALNHLKDEKLPLTAGICNPMIDMMLNPEFGNRRHLHSFLPN
MFLGAGMFIQFLIEQYIYFHSLLPEEIDFRISLYFVDYPLPFRVRSERSCYV
FYKLLHFEFVQNYAQRRIITNNFLRLELCYDSKLOAVQESVABAYAINRGWGSF
DDEMTGVNTINLGTNKAHLYPIILHLQVRNAPICVVGNGPSLDMLLPFTKANAHN
MIFGCTALKEPLKNAGIEPDQIETIERSYLDVLOSAPLGTLLCGNMVOPDALA
LAEAYIWRGSSASAYFGGASVVEPAPYVGNAGPALACLLSERVLICGLDCGYIK
GLSKRAQGYEREESTIPQNAVQNGSESEVIADALFSLSSAMMSKAIGIFPKPVV
INLGGAYIRGAPACTDFTLRNIDKSAIILKYSMAHQAQVLTDRDCRFQDSI
FYKEEILSLRESVGEISQVRNVAHKVLFKRIDMLYALILKSAKEPFGVLF
EGSISHILTMALHIPKODIALFYACQVEIIESGLNMTWSYKMLAMSHNSNL"
complement (4020..5078)
/locus_tag="HH0944"
complement (4020..5078)
/locus_tag="HH0944"
/notes="possible peptidase"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP77541.1"
/db_xref="GI:32262494"
/transl_table=11
/notes="WRNSSLPYLSLDSAOYVCGFSCDNALLVCIESRPFITDS
RYTLKELCKPHEVIESANFIOQVOELQKLIHKIVFNQELSVQMYEKLKNAL
KHKCEIIQENPHQKLRICKSEQIEGLIKTSQKLNKAPKFRGIYLNKIVKKSPEKE
LHYQAMQFLSHFGEISLSEFPPIYINANGAKPALPNKACYLAKNDILLFDAGIKYKR
YCSDMTRTAI KGEIHFQKQKPNKYQIYDVLKQEEAISKARSGMKAKQIDSL
ARDVIEKSGYGYFVHSTGCHGVGLDHELPRISLSECEVDGMSVSPGTYLPQEF
GVRIEDLVVMKNGRAEVL"
complement (5079..5555)
/locus_tag="HH0945"
complement (5079..5555)
/locus_tag="HH0945"
/codon_start=1
/transl_table=11
/product="3-dehydroquinase dehydratase II"
/protein_id="AAP77542.1"
/db_xref="GI:32262495"
/transl_table=11
/notes="WKILVIOQPNILNHLRPHFYIGNFTLDQIHNNLOAQAKONNAE
LEFFOSNPEGIIDLQECIGGEYAGVLINPAAFSHTSIAIADATASCGVPVIEVHIS
NIHAREEYRSKSYTCAVSAGVITGFAGFYHLALGILQIANEIALKAQQNNQNA"
complement (5675..7345)
/locus_tag="HH0946"
complement (5675..7345)
/locus_tag="HH0946"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP77543.1"
/db_xref="GI:32262496"
/transl_table=11
/notes="MTKKALYEEVIDEIGTQTSARFGLNIFNLVNFVSSYAYCAAA
VRATQDTPRAYDKHNDKYYGOTPEELDIGQNIKDSIFNTGNTYTTDLADIKKV
RDIRASGDITKLNTEDEKEDIMANYADEVKMDPEKGLNLRKNNTTDTIILDK
PENNVKTAQLKVIKSTGLNKDRYLEGDAADVDELKPPDDYKRHKHENLEKMIKEKGN
PENNOAVEKAQAKAEKLANNNCNRLMCENPRITAVATQSIVASGHIQAQFSDA
IIAALATLANGVIVWEVKDMFGSIDTETSLIKRIKRLKRTIEAFQATFGRAGFGAI
DAAVGVQIIFRSITAGKLKLVMDKRTALKSIYNGIVSVYIKGVEVSNRELIGIILKSL
FSAARVSTLAEGLKPLILASMAPPIVPFVFPFALLIAGAPAVVTHRSIDLVLTLF
GVFARADRAKURAEIADIIAQKQPALIEDREKLEELIASHKERLMSLDSFADYQK
AYQADDNGIYEGLNLCNLYDSSLISKMDGVDENILKPNRTKGLQ"
complement (7351..8502)
/locus_tag="HH0947"
complement (7351..8502)
/locus_tag="HH0947"
/notes="similar to Helicobacter pylori JHP0703; predicted
membrane-associated"
/codon_start=1

```

```

/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP77544.1"
/db_xref="GI:32262497"
/transl_table=11
/notes="MPLPLIPIILGGAALATAGYGVKKAIDAYDDSDQTAELYHERAKK
KFQSQARLDSKSAQVOKRSTGSLQEQIVTDELYADIVDKLKNADLQIDIVG
KEYTQGVANTQOSIVSLETTLGGITGGALAGALTGFMYGGVGLLASASTGTAIASLS
GVAATNATWALGGSLAAGSLMAGTMTVLGGIVAAPIAVMGSSVFAAKAEKKYDA
YAYDSVVADECEHQAQGLVWEQSLKRAKERSTILKESRVFVDASLEKIEDIMENKGY
TDIKNSNDVDEHRELMTQMLAEVIVAVINAPIMNDEDLTQELIEHQKCKBKMKREI
QKWQGMSEYSEKHEDLMKSFSGGKEYSENDRIIDGTGCFSGIPQGR"
complement (8561..8836)
/locus_tag="HH0948"
complement (8561..8836)
/locus_tag="HH0948"
/notes="similar to Helicobacter pylori HP0764"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP77545.1"
/db_xref="GI:32262498"
/transl_table=11
/notes="MMINSPPCTNCGMCKNIGGIKELKSPDLGNGICKYLDQTQSNLC
KIYQNRPDICRVDVMPFERYFKYTKAEAFYKLNQSKILQEKERID"
8937..9182
/locus_tag="HH0949"
8937..9182
/locus_tag="HH0949"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP77546.1"
/db_xref="GI:32262499"
/transl_table=11
/notes="MNKIQDEFIDNAYRIIILGNVKKARELKGLTQLELSLAIGHKSVS
VWSCAIYHKQHFNIHLLKIAQVLEVNQNPFEL"
complement (9329..11815)
/locus_tag="HH0950"
complement (9329..11815)
/locus_tag="HH0950"
/codon_start=1
/transl_table=11
/product="biotin sulfoxide reductase Bisc"
/protein_id="AAP77547.1"
/db_xref="GI:32262500"
/transl_table=11
/notes="MNVLSRQVLOAMGKGLAVPALVPFVGCDTESKKAIVESGVKSF
SENILVDGEVITGAHGVKAKIKDGKVISSNALKNVSNPLHSTIPDLIYSPTRIQ
YPMVRKSYLENPANRPELGADEWVRVSDEALALAKELKATYKQKGGVPAQSY

```

```

Query Match 34.6%; Score 166; DB 1; Length 302212;
Best Local Similarity 62.2%; Pred. No. 4.4e-33;
Matches 290; Conservative 0; Mismatches 155; Indels 21; Gaps 1;

QY 26 AATCTCAATTTTATAGACCATCTTACAAATTCACCTACCGGTATCCCATGCTTT 85
DB 233434 AACCCATAATGAATGTGGAGAAAATTCGTCAAATTTTACCACATCGTTATCGATGTTAC 233493

QY 86 TAGTGTAGATAATATAGAGTTACAAGCCAATAAAA-----AAA 124
DB 233494 TTGTTGATAGAGTTAGGAGCTTGTGCAAAATACAAGTATAGTGAACCAATGTTATA 233553

QY 125 TTGTCGCTTATAGAATATACATCTTTTAAATGAAGACGCTGTTTAAACGGGCATTTCCCTAATA 184
DB 233554 TCAAAACATATAAAATGTAACAAATTAATGAGAGGATTTTTCGGGCATTTTCCAAACA 233613

QY 185 AGCCCATTTTCCGGCGGTTTTCATCGTAGAGGCATGGCGCAACCGGAGGTTTTTTAG 244
DB 233614 AACCTATTATCCGGGTGTAATGCAGATTGAAGGAATGGCGCAACCGCGGTCTTTTAG 233673

QY 245 CCTTCACTAGCTGTGGGGTTTGAACCTGAATTCGCCCAACCAAAATCGGTATTTC 304
DB 233674 CATTTGTAAGTATGTTTGGTGAATAATGTTGACAGGCAAAAAATAAATTTGATATATTTA 233733

```

```
OY 305 TGACGATTGATAGGTTAAATTCCTCGATCCCTGTAAACCCAGCGACACGATTAGAAATACC 364
Db 233734 TGACGATAGATAAATGTTAAAGTTTCCTGTATCCCTGTCTGCCAGGTGATAGCTTGTATATG 233793
OY 365 ATTAGAGAGCTTAAAGCATAGGCGATGATCTGCGCAAGTGGTGGCAGCGCTCAAGTGG 424
Db 233794 AGCTCAAGGTTTAAACACAAAGGAAGTATATGGAACCTTGAGCAATGCTTTTGTG 233853
OY 425 ATGGCAAGAGTGTCTCCTGAAGCCGAATTAAGAGCCATGATTCGAGA 470
Db 233854 AGGATAAACTTGTAAGTGAAGCTGAACCTTAAGCTATGATTACTGA 233899

RESULT 7
BX571657/c 347961 bp DNA linear BCT 17-APR-2005
LOCUS Wolinella succinogenes, complete genome; segment 1/7.
DEFINITION BX571657 BX571656
ACCESSION BX571657.1 GI:34482172
VERSION complete genome.
KEYWORDS Wolinella succinogenes
SOURCE Wolinella succinogenes
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Wolinella.
REFERENCE
1 Baar,C., Eppinger,M., Raddatz,G., Simon,J., Lanz,C., Klimmek,O.,
Nandakumar,R., Gross,R., Rosinus,A., Keller,H., Jagtap,P.,
Linke,B., Meyer,F., Lederer,H. and Schuster,S.C.
Complete genome sequence and analysis of Wolinella succinogenes
Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11690-11695 (2003)
14500908
2 Schuster,S.C.
Direct Submission
Submitted (15-MAY-2003) Max-Planck Institut for Developmental
Biology, Spemannstr. 35, 72076 Tuebingen, GERMANY
LOCATION/Qualifiers
1. 347961
/organism="Wolinella succinogenes"
/mol_type="genomic DNA"
/strain="DSMZ 1740"
/db_xref="taxon:844"
1. .1314
/gene="dnaA"
/locus_tag="WS0000"
1. .1314
/gene="dnaA"
/locus_tag="WS0000"
/locus_start=1
/codon_start=1
/transl_table=11
/product="dnaA, probable chromosomal replication initiator
protein"
/protein_id="CAE09174.1"
/db_xref="GI:34482173"
/db_xref="GOA:Q7MSY2"
/db_xref="InterPro:IPR001957"
/db_xref="UniProt/Swiss-Prot:Q7MSY2"
/translation="MLGDTTLKQLKDEISTLEYERVIKQLRVDEASRTDLAVIAPN
QPIANWVKTKYGERLAHLFELSTGIRPKIEIRLGSCLKQVSKSSPKAGSKGQKSTIL
NPSFTFDSFVGNRNFAVEVSONAKQCIAYNPILLIYGGTGLGTHLINSIGNYV
AKGSVIYVTSQFLNDFLYHIRNNTMDRFRDKYACDYLIDVDFGCKPQIQEAF
FHTFELHNKQIVLTSKTPQIAGLEERLKSREMGVSDIQPELETKNIILKK
KECFDGIYLSNEIISYIATNMNNINREIGIILKNAYANLHMQEITLQPAKNVLKEQ
IKBERNITLSENIIEVVSKELVNPKPAEIKSGSKNIAARRIVIFLARTLTNPSMPS
LAQFFGMDKDHSSVKAMKTVKTEIEDANFQWVIEELKNKIKRNN"
1474. .2541
/gene="DNAN"
/locus_tag="WS0001"
1474. .2541
/gene="DNAN"
/locus_tag="WS0001"
/EC_number="2.7.7"
/codon_start=1
```

```
/transl_table=11
/product="DNA POLYMERASE III, BETA CHAIN"
/protein_id="CAE09175.1"
/db_xref="GI:34482174"
/db_xref="GOA:Q7MAV2"
/db_xref="InterPro:IPR001001"
/db_xref="UniProt/TREMBL:Q7MAV2"
/translation="MNIVPKSSLNLTNLQPFLEKKASQITSHVYFKAQGSSELPF
QATDYIGLLAKTESVQIQEGIAVTNGKKILDIVRLEKEGITLQOEYLIHKQEK
SFKLPMFAEEPPSPSEELPKIEINSMHLVQSMKKILPAIDNNNPKFELNGALLD
IKESFNVAITDTRRLAVIKYENPINKLALIVPKAIVRIQKLFEEVELSHGETHL
ILKSNYLFPTKLISGKYDPYERLIIPKETHNTLPEKILIEAKLINSLSNDIKVTF
TKEEILFESLSENSEAKTQVEVETGITLTHLTGINSKYLDDVLGQIDTSEFKIGINE
SNTPFILKSENFSTVIMPVM"
2561. .4879
/gene="gyrB"
/locus_tag="WS0002"
2561. .4879
/gene="gyrB"
/locus_tag="WS0002"
/BC_number="5.99.1.3"
/codon_start=1
/transl_table=11
/product="DNA GYRASE SUBUNIT B"
/protein_id="CAE09176.1"
/db_xref="GI:34482175"
/db_xref="GOA:Q7MAV1"
/db_xref="InterPro:IPR001241"
/db_xref="InterPro:IPR002288"
/db_xref="InterPro:IPR003594"
/db_xref="InterPro:IPR006171"
/db_xref="InterPro:IPR011557"
/db_xref="InterPro:IPR011558"
/db_xref="InterPro/TREMBL:Q7MAV1"
/translation="MQNQYSDANIKVLKGLAVRKEPKGMVIGDTNVNGLHLIVEV
DSNIDEAMAGHCNTISITLTKESAIIEDNGRGIIPDMHPTENIPAAITVLTVLHAGG
KFDKDSYKSGGLHGVSVNALSCKLVMTRKNGNVYTFQFSGCPIKGLDILIGTS
KGLRQTFIEITDDIEFVTVFDRLEARRFELAYLNNDYDEKVLSPVNNIKTPEGG
GGLRQFVEHLNKKVLVSQVFSFSDNESDTEVEIALYNDGYDEKVLSPVNNIKTPEGG
THEAGFRAGLSRAITSVIDANANAREKDNKITGDVREGLVAIVSVKMEPQFEGQTK
GKLGSFVRPIVQKLTYEKIAKFEENPNDAKAIMQKALMAAREAKARELTKK
DSLMVTLPKGLADCKSPISIELYLVEGDSAGSASQGRDVRVQAILPLKGLILNV
EKSRLDKLSDEIKNMITALGCGIEEFIDNIRIYQKIIMTDADVDSGSHIQTLILT
FFRPLKPVIEAGLYIAQPLYRFKKGKEIYLRDEKALSEYLIENGIENFAFEGVG
TPDLLEIRPASHVRLTLEKEKRFAMIDVVKVMIHRDLTSLPMSELYIKLEBEVQQ
KGYNILNKVVEEKIHLVQTHEGLVDLVDDSLFTHPLPEAYVYVHNLILERDLSPL
GENDILTLBEEESAKGAYIQRYKGLGEMNPQLWETTTMPENRLLKVKQITDDME
ASDFSLFMGDEVEPRRAYIQDHAKDVKKHLDV"
4890. .5264
/locus_tag="WS0003"
4890. .5264
/locus_tag="WS0003"
/notes="HYPOTHETICAL PROTEIN CUI1724C"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAE09177.1"
/db_xref="GI:34482176"
/db_xref="GOA:Q7MSY1"
/db_xref="InterPro:IPR001474"
/db_xref="UniProt/TREMBL:Q7MSY1"
/translation="MRYGEMVKNFPEEIEVWPNRNDRHVTIKITLPEFSCLCPRSG
YPIYATVYIEYVSSLVNVELKAIKLYINSFDRHVSHEDSANEIYDLLYKKLSPELY
LRMDFNPRGNVHTIIIEIDSKKN"
5274. .5900
/locus_tag="WS0004"
5274. .5900
/locus_tag="WS0004"
/codon_start=1
/transl_table=11
/product="PUTATIVE INTEGRAL MEMBRANE PROTEIN"
/protein_id="CAE09178.1"
/db_xref="GI:34482177"
```



```

/misc_feature
1. .1323
/gene="dnaA"
/gene="dnaA"
/notes="Cj0001, dnaA, probable chromosomal replication
initiator protein, len: 440 aa; similar to many e.g.
DNAA_ECOLI (467 aa); fasta scores; opt: 839 z-score: 949.6
E(): 0, 32.3% identity in 470 aa overlap. 39.9% identity
to HP1529. Contains PS0017 ATP/GTP-binding site motif A
(P-loop) and Pfam match to entry PF00308 bac_dnaA,
Bacterial dnaA protein, score 419.90, E-value 4e-141"
/codon_start=1
/transl_table=11
/product="chromosomal replication initiator protein"
/protein_id="CAB72494.1"
/db_xref="GI:6967506"
/db_xref="GOA:Q9PJB0"
/db_xref="InterPro:IPR001957"
/db_xref="InterPro:IPR003593"
/db_xref="InterPro:IPR010921"
/db_xref="UniProt/Swiss-Prot:Q9PJB0"
/translacion="WNPSQILENLKKEISEYENYSNLKFEKQSKADLLVFNAPN
ELMAKFIOTYKGGKIAHFYVQSGNKAIINIOAQAKQSNKSTKIDIAHKAQSTILN
PSFTESFVGDGSKYVACACKAIAHKDKLGLNPIFVVGPTGLGKTHLLQAVGNAS
LEMGGKIYATSENFINTPSNLKNGSLDKFHEKVRNCVLLIDVQFLGDKTDIOEE
FFFIFNEIKNDQIMTSDNPPNMLKGITERLKRFAHGLIADITPPOLDTKIALIR
KKCFENDINLNDIIINYATSGDNIREEGIIISLNAYATILGQETILELAKSVMKD
HIKEKNITIDILSLVCKEFNIXPSDVSKNKTQNIVTARRIVILARALTALTMP
OLANYFEMKDHTAISHNVKKITEMIENDASLAKAKIEELKNILVKSQS"
301. .1239
/gene="dnaA"
/notes="Pfam match to entry PF00308 bac_dnaA, Bacterial
dnaA protein, score 419.90, E-value 4e-141"
427. .450
/gene="dnaA"
/notes="PS00017 ATP/GTP-binding site motif A (P-loop)"
1483. .2550
/gene="dnaA"
1483. .2550
/gene="dnaA"
/EC_number="2.7.7.7"
/notes="Cj0002, dnaN, probable DNA polymerase III, beta
chain, len: 355 aa; similar to e.g. DP3B_ECOLI DNA
polymerase III, beta chain (EC 2.7.7.7) (366 aa), fasta
scores; opt: 349 z-score: 372.3 E(): 1.7e-13, 19.6%
identity in 367 aa overlap. 34.2% identity to HP0500.
Contains Pfam match to entry PF00712 DNA_pol3_beta, DNA
polymerase III beta subunit, score 119.80, E-value
5.2e-32"
/codon_start=1
/transl_table=11
/product="DNA polymerase III, beta chain"
/protein_id="CAB72495.1"
/db_xref="GI:6967507"
/db_xref="GOA:Q9PJA9"
/db_xref="InterPro:IPR001001"
/db_xref="UniProt/TrEMBL:Q9PJA9"
/translacion="MKLSINKNTLESAVILCNAYVVEKSDSTTJSHLFFHADEKLLI
KASDVEIGINYKIKRIVSSGFTANAKSIADVIKSLNNEEVLSTIDNPLFVRKQS
TKYKLPFHEDFPFNPTEGKNQFDISDSLSRSLKPLPSIDTNPKYSLNGAFID
IKTDKINFPTKRLAIYITLEKANNQSFSPKPAIMEMQKLFYEKIEIFDYDNML
IAKNENFEFTKLINDIPDYEKVIPKTPKQLSFSTEDFIDSLKISVVTWKRLHF
NKDKIIIFEGISLDNMEAKTELEITQTVGSEEFNLTIKIKLLDLPLTSIEBEKFTLSVNE
PNSAFIVKSGLSMIIMPIL"
1483. .2535
/gene="dnaN"
/notes="Pfam match to entry PF00712 DNA_pol3_beta, DNA
polymerase III beta subunit, score 119.80, E-value
5.2e-32"
2579. .4888
/gene="gyrB"
2579. .4888
/gene="gyrB"
/EC_number="5.99.1.3"

misc_feature
/notes="Cj0003, gyrB, probable DNA gyrase subunit B, len:
769 aa; similar to many e.g. GYRB_BACSU DNA gyrase subunit
B (EC 5.99.1.3) (638 aa), fasta scores; opt: 2130 z-score:
3165.8 E(): 0, 47.7% identity in 773 aa overlap. 61.2%
identity to HP0501. Contains PS00177 DNA topoisomerase II
signature and Pfam matches to entry PF00204 DNA_topoisoiI,
DNA topoisomerase II (N-terminal region), score 915.20,
E-value 1.9e-271, and to entry PF00986 DNA_gyraseB_C, DNA
gyrase B subunit, carboxyl terminus, score 154.00, E-value
2.6e-42"
/codon_start=1
/transl_table=11
/product="DNA gyrase subunit B"
/protein_id="CAB72496.1"
/db_xref="GI:6967508"
/db_xref="GOA:O87667"
/db_xref="InterPro:IPR000565"
/db_xref="InterPro:IPR001241"
/db_xref="InterPro:IPR002288"
/db_xref="InterPro:IPR003594"
/db_xref="InterPro:IPR006171"
/db_xref="UniProt/Swiss-Prot:O87667"
/translacion="MOENYCASNLKVLKLEAVRKPQMYIGDNTNIGLHHMIVEVD
NSIDEAMAGCHDTIDVEITTEGCIIVSDNGRGIIVDMHPTENMPTLTVLTVLHAGGK
PKDITYKSGGLKGVSVVVALSKLVATVERNGEIQREFSGKVISFEGVIGKSK
KGTTFIEFPDDQIFEVTEFDYELAKRFLAYLNPKITINFKNRVGKHESHFPEG
GISQFVTDNLKEALTKAIFPSVDEEDVNVVALLNYDTSYENLLSFVNNIKTPDGGT
HAGPFMGLTRVTSNIETANASAKDNKIKTDGDRGLIAIVSVKPEPQEQTKG
KLGSYVRPIVSKASFYELTYPEENPIEAKIMWKAARGEAAKARELRYKKE
SLSVGTLPGLADQCKSDPSESEIYLVGDSAGSAGKAGRSFQAILPLRGKILNVE
KARKDKILKSEQIQNMITARGCGIGEDFDLSKLRYHKIIMTDADVDSHQTLLLTTF
PFRFMELVANGHIYLAQPPLYKKAKKAIYLDKDEKALSELYLTETGHEGLNYGIGM
NDLKDYIKIVAYRAILKDLERKNFVSVIRYMIENSLVKNNEELFSVKQLETOQ
GHNILNHYINENEIRAFVQONGLELVINEELFTHPLYERASYIFDKIKRDSLEFDK
DILVELDEVETNAKKGATIQRYKGLGEMNPQLWETTMDSVSRLLKLTIEDAQSAND
TFNLFMGDEVEPRDRIQAHAKVDKHLV"
2606. .4168
/gene="gyrB"
/notes="Pfam match to entry PF00204 DNA_topoisoiI, DNA
topoisomerase II (N-terminal region), score 915.20,
E-value 1.9e-271"
3830. .3856
/gene="gyrB"
/notes="PS00177 DNA topoisomerase II signature"
4649. .4852
/gene="gyrB"
/notes="Pfam match to entry PF00986 DNA_gyraseB_C, DNA
gyrase B subunit, carboxyl terminus, score 154.00, E-value
2.6e-42"
complement(4916. .5257)
/gene="Cj0004c"
complement(4916. .5257)
/gene="Cj0004c"
/notes="Cj0004c, possible periplasmic protein, len: 113 aa.
No Hp match. Contains possible N-terminal signal sequence"
/codon_start=1
/transl_table=11
/product="putative periplasmic protein"
/protein_id="CAB72497.1"
/db_xref="GI:6967509"
/db_xref="GOA:Q9PJA7"
/db_xref="UniProt/TrEMBL:Q9PJA7"
/translacion="MKKIIILALFLSASWAQNLINPDITGLIIDPDSPLVANCLAC
HGSNLIYNMHASRKAWLAARWMQDSEGLWIEPEDEKILNLEYKGYEKYDTRRRI
PLAILLQNKTH"
complement(5260. .6498)
/gene="Cj0005c"
complement(5260. .6498)
/notes="Cj0005c, possible molybdenum containing
oxidoreductase, len: 412 aa; similar to many eukaryotic
oxidoreductases e.g. SUOX_HUMAN sulfite oxidase precursor
(488 aa), fasta scores; opt: 543 z-score: 338.2 E():

```

```
1.4e-11, 29.8% identity in 372 aa overlap and NIA_PETHY
nitrate reductase from Petunia hybrida (909 aa), fasta
scores; opt: 281 z-score: 312.5 E(): 3.8e-10, 28.9%
identity in 395 aa overlap. No Hp match. Contains Pfam
match to entry PF00174 oxidoredo molyb, Oxidoreductase
molybdopterin binding domain, score 45.20, E-value 1e-12"
/codon_start=1
/transl_table=11
/products="putative molybdenum containing oxidoreductase"
/protein_id="CA872498.1"
/db_xref="GI:6967510"
/db_xref="GOA:Q9PJA6"
/db_xref="InterPro:IPR000572"
/db_xref="InterPro:IPR005066"
/db_xref="InterPro:IPR008335"
/db_xref="UniProt/TrEMBL:Q9PJA6"
/translation="MKQNDQKRNDRFLKNIIGLFLGISVLNFSNPFSLGSKALAKE

Query Match      28.2%; Score 135.2; DB 1; Length 314150;
Best Local Similarity 58.4%; Pred. No. 6.3e-25;
Matches 258; Conservative 0; Mismatches 178; Indels 6; Gaps 1;

QY 39 TATAGAGCATATCTTACAAATCTACCTCACCGCTATCCCATGCTTTTACTGGATGAAT 98
Db 250598 TGTATGCAAAATCAAGAAATCTTACACACCGTTATCTTTTACTAGTAGATAAAT 250657

QY 99 TATAGAGTTACAGCCCAATAAAAAATTCGCTTATAAGAAATATCACTTTTAAATGAAGA 158
Db 250658 TACAGAAATTAAGATTAAAGAGTGGTCTAGGATATAAAAAACATCAGTATAGTGACCA 250717

QY 159 CGTGTTTAAGGGCATTTCCCTAATAAGCCCATTTTCCGGGGCTTTTGATCGTAGAGGG 218
Db 250718 TGTGTTTTATGGGACATTTTCCAGGGCATCTTATTTATCTCGAGTTTGTGATTTTGAAGG 250777

QY 219 CATGGCGCAACGGGAGGTTTGTAGCTTCACTAGCTTTGTGGGGTTTGACCTGAAT 278
Db 250778 TATGGCTCAACAGGTGGAGTTTGTAGCTTTTGAAGTAT-----GGAAGATAAGTGGA 250831

QY 279 CGCCAAAACAAAAATCGTGTATTTCATGACGATTGATAAGTTTAAATTCGCGATCCCTGT 338
Db 250832 TCCAAAAGTAAGTAGTTTATTTCACAGGCATAGATGAGCAAAATTTAGAATCTGT 250891

QY 339 AACCCAGCGCAGAGATTAGATACCAATTTAGAAGTCTTTAAAGCATAGGGCGATGATCG 398
Db 250892 GCGTCTCGAGATAGGCTTGATTATGAATGAGCGTGGTTAAATAATCGTGAATATGTG 250951

QY 399 GCAAGTGGTGCGCAGCGCTCAAGTGTGCAAGTGGTTCGCTGAAGCCGAATTGAAGC 458
Db 250952 GATTTTAAAGGGCAAGCTTTTGTAGATGGAATTTAGTTGCAGAGGCCGAGCTTAAAGC 251011

QY 459 CATGATTGCAGAGAGAGATTAA 480
Db 251012 TATGATAGTGGATAAATAATGA 251033
```

RESULT 9
CP000025_02
WFCOMENT

Sequence split into 18 fragments LOCUS CP000025 Accession CP000025

Fragment Name	Begin	End
CP000025_00	1	110000
CP000025_01	100001	210000
CP000025_02	200001	310000
CP000025_03	300001	410000
CP000025_04	400001	510000
CP000025_05	500001	610000
CP000025_06	600001	710000
CP000025_07	700001	810000
CP000025_08	800001	910000
CP000025_09	900001	1010000
CP000025_10	1000001	1110000
CP000025_11	1100001	1210000
CP000025_12	1200001	1310000

CP000025_13 1300001 1410000
CP000025_14 1400001 1510000
CP000025_15 1500001 1610000
CP000025_16 1600001 1710000
CP000025_17 1700001 177831
Continuation (3 of 18) of CP000025 from base 200001 (CP000025 Campylobacter jejuni RM1221)

Query Match 27.8%; Score 133.6; DB 1; Length 110000;
Best Local Similarity 58.1%; Pred. No. 1.8e-24;
Matches 257; Conservative 0; Mismatches 179; Indels 6; Gaps 1;

QY 39 TATAGAGCATATCTTACAAATCTTACCTCACCGCTATCCCATGCTTTTACTGGATGAAT 98
Db 86911 TGTATGCAAAATCAAGAAATCTTACACACCGTTATCTTTTACTGGTAGATAAAT 86970

QY 99 TATAGAGTTACAGCCCAATAAAAAATTCGCTTATAAGAAATATCACTTTTAAATGAAGA 158
Db 86971 TACAGAAATTAAGATTAAAGAGTGGTCTAGGATATAAAAAACATCAGTATAAGTGATCA 87030

QY 159 CGTGTTTAAGGGCATTTCCCTAATAAGCCCATTTTCCGGGGCTTTTGATCGTAGAGGG 218
Db 87031 TGTGTTTTATGGGACATTTTCCAGGGCATCTTATTTATCTCGAGTTTGTGATTTTGAAGG 87090

QY 219 CATGGCGCAACGGGAGGTTTGTAGCTTCACTAGCTTTGTGGGGTTTGACCTGAAT 278
Db 87091 TATGGCTCAACAGGTGGAGTTTGTAGCTTTTGAAGTAT-----GGAAGATAAAGTGGA 87144

QY 279 CGCCAAAACAAAAATCGTGTATTTCATGACGATTGATAAGTTTAAATTCGCGATCCCTGT 338
Db 87145 TCCAAAAGTAAGTAGTTTATTTCACAGGCATAGATGAGCAAAATTTAGAATCTGT 87204

QY 339 AACCCAGCGCAGAGATTAGAAATACCAATTTAGAAGTCTTTAAAGCATAGGGCGATGATCG 398
Db 87205 GCGTCTCGAGATAGGCTTGATTATGAATCAGCGTGGTTAAATCTGGTGAATATGTG 87264

QY 399 GCAAGTGGTGCGCAGCGCTCAAGTGTGCAAGTGGTTCGCTGAAGCCGAATTGAAGC 458
Db 87265 GATTTTAAAGGGCAAGCTTTTGTAGATGGAATTTAGTTGCAGAGGCCGAGCTTAAAGC 87324

QY 459 CATGATTGCAGAGAGAGATTAA 480
Db 87325 TATGATAGTGGATAAATAATGA 87346

RESULT 10
AE015936
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AE015936 300029 bp DNA linear BCT 06-FEB-2003
Clostridium tetani E88, section 1 of 10 of the complete genome.
AE015936 AE015927
AE015936.1 GI:28202269
Clostridium tetani E88
Clostridium tetani E88
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.

REFERENCE
1 (bases 1 to 300029)
Bruggemann,H., Baumer,S., Fricke,W.F., Wierzer,A., Liesegang,H.,
Decker,I., Herzberg,C., Martinez-Arias,R., Merkl,R., Henne,A. and
Gottschalk,G.
The genome sequence of Clostridium tetani, the causative agent of
tetanus disease
Proc. Natl. Acad. Sci. U.S.A. 100 (3), 1316-1321 (2003)
12552129

REFERENCE
2 (bases 1 to 300029)
Bruggemann,H., Baumer,S., Fricke,W.F., Wierzer,A., Liesegang,H.,
Decker,I., Herzberg,C., Martinez-Arias,R., Merkl,R., Henne,A. and
Gottschalk,G.
Direct Submission
Submitted (31-OCT-2002) Goettingen Genomics Laboratory, Institute
of Microbiology and Genetics, Georg-August University,
Grisebachstr. 8, Goettingen 37077, Germany
Location/Qualifiers
1 .300029

gene	<pre>/organism="Clostridium tetani E88" /mol_type="genomic DNA" /strain="Massachusetts" /sub_strain="E88" /db_xref="taxon:212717" complement(1..2304) /gene="mecB" /locus_tag="CTC00060" complement(1..2304) /gene="mecB" /locus_tag="CTC00060" /codon_start=1 /transl_table=11 /product="negative regulator of genetic competence mecB" /protein_id="AAO34716.1" /db_xref="GI:28202270" /translacion="MKLSCVNKNVATITFTTFETGKPOLGVCACAKKNNIOPLNO IIEAGVTPDEMITYAEQNKLLDNMDINELLNDFNMMNMMNGMMFKVQDDFID KDKNEENTNSRNDKSGNSNVKTRVPKPKNRRLDKYAINLNDKADKKIDVVVGR DKIEIRIILNRRNNPNLIGEPGVGKTAIEGLAVKIIIEKNVPYIKLFDKBIYLLD LTSVAGTQFQGEGRMAKIVDEVTESENVILVIDEHNIIIGAGEAQGALNAANIL KPALARGIOVIGATLEEKRYIEKDSALERFPQILVEBPSVKETIIEILNNKIYY EDYHKVTKSKIIIEAVKLSRYITDRYLPDKAIDVIDEASRVLNRNKNLVEFKIAK BELGLQNMNLNEANGNYEKAAQYKTEYKLNKKNVBELEKSKVELSVDDIAYVIEA WTKIPQVQRKEAEKLNLEORLHKRLIGQDKAVRKLSTIRRNRLGFKSKKRPASF IFVGTGVGTTELVSRLACELFGDDESLRIDMSYMEKHTASKLIGSPGPGVGYDEG QQLTEKIRKPYSVLLDEIEKAHPDIENMLLQVLELDGRLTDNQGRTVFDNSVIIMT SNLGTNFKESSIGFNTWSIOIENSIKDALKETFKPEFNRIIDEIVIPEKLNKEBLYD IIDIMEVSDLSSEKNIKVHPTKEVKQPIKLVGVDEKYGARELKRAIQRYIEBSE QYFNEVKEGNTILVDIYEESIYIKNM" 2487..3923 /locus_tag="CTC00061" 2487..3923 /locus_tag="CTC00061" /EC_number="2.7.8.-" /codon_start=1 /transl_table=11 /product="putative cardiolipin synthetase 2" /protein_id="AAO34717.1" /db_xref="GI:28202271" /translacion="WKISFTSLSPFFFINIIITLVIILERKRPEKTIAMLIITLL PPGILFVFLGRNWKHLNNDPSPKIRELIYRSMKDIDDPPYVPLIHIAENSESP LFLNNSIEIFDGTGKFAKREILNAKHIIHLEYIVKNDIIGNEIKNLLIEKSLBG VEVREILDRVGSIKLGEPFINDIMXNGIDVOQVSYPFLAPLLKKNITQVNNRHRKLI IDGVEFTGGINIGDEVYLGKGLGYWRDTHIMVKGDFVLGLOAVFIDDFWTIKXAND SILINTDHSYFPKPNQSGKKLMOLAKSGPNSPYSIMQGLIKMLMYATMDHLYITPY FVPSSEIMALKIACLGIDRILFPQGDHFTVTSASRTYLAELLKSKAKVHFYDEN SFVHAKIITVDGKVTIGTANNDIRSDFLNBIINAIYDEETITKLEEMFFEDLNHSK NVTGQDNFETSNFMKFIETALTRFSSLL" 4027..4548 /locus_tag="CTC00062" 4027..4548 /locus_tag="CTC00062" /codon_start=1 /transl_table=11 /product="DNA polymerase, bacteriophage-type" /protein_id="AAO34718.1" /db_xref="GI:28202272" /translacion="WMSGGIPCDILIFIGBAPGKTEVEBGKPFVGAAGKNFQYVLDIN LNKDIRITNTCYFRPKIKMSGSGRTISNRPPTKEINLFOHLKKEIETVNPKEII ITLGTNPLKNTFQNVNATIGHRHGSLIFSEKLNHIPPVHPHSALINRDTSPKESYKE DWCKIADILKGPK" complement(4584..5144) /locus_tag="CTC00064" complement(4584..5144) /locus_tag="CTC00064" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAO34719.1" /db_xref="GI:28202273" /translacion="MYDITYTIERGINMSIKNSINIGLSKIATKYTGDNATVAQKSS EFVEISKLNINISSEKNNTIEELYREIGRVVYSKYENGEIIDKDLEDSQQOLGSEYINKI</pre>
CDS	<pre>/organism="Clostridium tetani E88" /mol_type="genomic DNA" /strain="Massachusetts" /sub_strain="E88" /db_xref="taxon:212717" complement(1..2304) /gene="mecB" /locus_tag="CTC00060" complement(1..2304) /gene="mecB" /locus_tag="CTC00060" /codon_start=1 /transl_table=11 /product="negative regulator of genetic competence mecB" /protein_id="AAO34716.1" /db_xref="GI:28202270" /translacion="MKLSCVNKNVATITFTTFETGKPOLGVCACAKKNNIOPLNO IIEAGVTPDEMITYAEQNKLLDNMDINELLNDFNMMNMMNGMMFKVQDDFID KDKNEENTNSRNDKSGNSNVKTRVPKPKNRRLDKYAINLNDKADKKIDVVVGR DKIEIRIILNRRNNPNLIGEPGVGKTAIEGLAVKIIIEKNVPYIKLFDKBIYLLD LTSVAGTQFQGEGRMAKIVDEVTESENVILVIDEHNIIIGAGEAQGALNAANIL KPALARGIOVIGATLEEKRYIEKDSALERFPQILVEBPSVKETIIEILNNKIYY EDYHKVTKSKIIIEAVKLSRYITDRYLPDKAIDVIDEASRVLNRNKNLVEFKIAK BELGLQNMNLNEANGNYEKAAQYKTEYKLNKKNVBELEKSKVELSVDDIAYVIEA WTKIPQVQRKEAEKLNLEORLHKRLIGQDKAVRKLSTIRRNRLGFKSKKRPASF IFVGTGVGTTELVSRLACELFGDDESLRIDMSYMEKHTASKLIGSPGPGVGYDEG QQLTEKIRKPYSVLLDEIEKAHPDIENMLLQVLELDGRLTDNQGRTVFDNSVIIMT SNLGTNFKESSIGFNTWSIOIENSIKDALKETFKPEFNRIIDEIVIPEKLNKEBLYD IIDIMEVSDLSSEKNIKVHPTKEVKQPIKLVGVDEKYGARELKRAIQRYIEBSE QYFNEVKEGNTILVDIYEESIYIKNM" 2487..3923 /locus_tag="CTC00061" 2487..3923 /locus_tag="CTC00061" /EC_number="2.7.8.-" /codon_start=1 /transl_table=11 /product="putative cardiolipin synthetase 2" /protein_id="AAO34717.1" /db_xref="GI:28202271" /translacion="WKISFTSLSPFFFINIIITLVIILERKRPEKTIAMLIITLL PPGILFVFLGRNWKHLNNDPSPKIRELIYRSMKDIDDPPYVPLIHIAENSESP LFLNNSIEIFDGTGKFAKREILNAKHIIHLEYIVKNDIIGNEIKNLLIEKSLBG VEVREILDRVGSIKLGEPFINDIMXNGIDVOQVSYPFLAPLLKKNITQVNNRHRKLI IDGVEFTGGINIGDEVYLGKGLGYWRDTHIMVKGDFVLGLOAVFIDDFWTIKXAND SILINTDHSYFPKPNQSGKKLMOLAKSGPNSPYSIMQGLIKMLMYATMDHLYITPY FVPSSEIMALKIACLGIDRILFPQGDHFTVTSASRTYLAELLKSKAKVHFYDEN SFVHAKIITVDGKVTIGTANNDIRSDFLNBIINAIYDEETITKLEEMFFEDLNHSK NVTGQDNFETSNFMKFIETALTRFSSLL" 4027..4548 /locus_tag="CTC00062" 4027..4548 /locus_tag="CTC00062" /codon_start=1 /transl_table=11 /product="DNA polymerase, bacteriophage-type" /protein_id="AAO34718.1" /db_xref="GI:28202272" /translacion="WMSGGIPCDILIFIGBAPGKTEVEBGKPFVGAAGKNFQYVLDIN LNKDIRITNTCYFRPKIKMSGSGRTISNRPPTKEINLFOHLKKEIETVNPKEII ITLGTNPLKNTFQNVNATIGHRHGSLIFSEKLNHIPPVHPHSALINRDTSPKESYKE DWCKIADILKGPK" complement(4584..5144) /locus_tag="CTC00064" complement(4584..5144) /locus_tag="CTC00064" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAO34719.1" /db_xref="GI:28202273" /translacion="MYDITYTIERGINMSIKNSINIGLSKIATKYTGDNATVAQKSS EFVEISKLNINISSEKNNTIEELYREIGRVVYSKYENGEIIDKDLEDSQQOLGSEYINKI</pre>
gene	<pre>IBEDKIDNIKSTKVNKKVKCDTDEDICTEYKSTEDSKIEDAKVSELHDSILKPD DELVNECEGPGCNFTWTDGRDDDLIK" complement(5249..5324) /product="trNA-Lys" /notes="codon recognized: AAA" complement(5340..5427) /product="5S ribosomal RNA" complement(5606..8450) /product="23S ribosomal RNA" 5938..6594 /locus_tag="CTC00065" 5938..6594 /locus_tag="CTC00065" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAO34720.1" /db_xref="GI:28202274" /translacion="MRRADIEVNPLPVDVDSWGRSACYPRGSFYPLSDGPPTRNHRIT KPDFPCSTCMCRSQAPLCLYLIRAI SDRABGTGRLRYFLGGDRPSPQTAHLTWSRDQ IHGPRLEFOYQGGIPRMPQKLTLLSLPILYRQYRNSMLSYKALRGLSLSRV ASIFCTTISPDLLLRQCPSHYAIRAGRNLPDKEFRYLRTVITVTAAGWGLSSPLRLR " complement(8715..10223) /product="16S ribosomal RNA" 10753..11397 /locus_tag="CTC00066" 10753..11397 /locus_tag="CTC00066" /codon_start=1 /transl_table=11 /product="conserved protein" /protein_id="AAO34721.1" /db_xref="GI:28202275" /translacion="MKNKKIKLITFIIVIFAYNVKLVNDKINSLPMNDKSIKQMSR KITRDAVVKIQEKQDLITMEVBLNKNKVIDLDSWGNLDKFKKVIISDFYGTGIYTD LAKLNRENVLIQSKTITIRIPKPKVKSITLNEKDTFTKTENGUFRGEEKITPAESQ ILSKAKERNQELNKHVESALSCTKSIEKSIQSIILLPYENINIVIKFIDG" complement(12432..13208) /locus_tag="CTC00067" complement(12432..13208) /locus_tag="CTC00067" /codon_start=1 /transl_table=11 /product="putative polyferredoxin" /protein_id="AAO34722.1" /db_xref="GI:28202276" /translacion="MRGVLCYFGTGNTKMIADKPKYFNMKNIELDIINIIEEKQDME LSEYNFLVVGTPYIGELPKIVDTFLNLPNNGKNIKITILYSTQGTAKASALNVIESKI ENKNYTVSGKIMFKMTNNFYISGKASKEIIEILISCERKVNNTVNNFIEGKYIKE KVNKIRVTLSKIASKHFNHYLNLSKNLSSSENCIRGCLCLRNCPQNNITPENGHVFI HSCKMCLRCIYICPTNEIRYKKNKIDQIKRIIKSLKIK" 13557..14816 /locus_tag="CTC00068" 13557..14816 /locus_tag="CTC00068" /codon_start=1 /transl_table=11 /product="conserved protein" /protein_id="AAO34723.1" /db_xref="GI:28202277" /translacion="MKNLSIYNVIRLLLYFACFIIFTDTPISIFTLKNNIILQPD ILITDYLEISIGIGASFVNSATLCIIVILIRLKLVSIGSVAGLFTVCGFSLFGKQL LNVWPLFLGTWLYSKYQKEPPSKYILVALFGSTLSPAVSQLSFSGHFETSGLIIGIS LGIPLMGFILPISISYCLNHEGYNLNSGPAAGLVGTFPMSILRSFGINFORLWST KYTVLLIIFLPSILFLSMILVFCNLDPSFNKLDKIKHSGKPNKNTIPVWLGTVIAGTINI MGILGIFMWYIILIVDGLNGPIAGTIVTIVGAFGKGFNTIPVWLGTVIAGTINI WSINSPMLLATLFTSTLAPISGVFGPIYGIAGFLHTCIWMNIGHLHAGLNLYNNGF SGGFVALIILPIITNLKVDTONELQI" complement(14871..15398) /locus_tag="CTC00069" complement(14871..15398) /locus_tag="CTC00069"</pre>
CDS	<pre>IBEDKIDNIKSTKVNKKVKCDTDEDICTEYKSTEDSKIEDAKVSELHDSILKPD DELVNECEGPGCNFTWTDGRDDDLIK" complement(5249..5324) /product="trNA-Lys" /notes="codon recognized: AAA" complement(5340..5427) /product="5S ribosomal RNA" complement(5606..8450) /product="23S ribosomal RNA" 5938..6594 /locus_tag="CTC00065" 5938..6594 /locus_tag="CTC00065" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAO34720.1" /db_xref="GI:28202274" /translacion="MRRADIEVNPLPVDVDSWGRSACYPRGSFYPLSDGPPTRNHRIT KPDFPCSTCMCRSQAPLCLYLIRAI SDRABGTGRLRYFLGGDRPSPQTAHLTWSRDQ IHGPRLEFOYQGGIPRMPQKLTLLSLPILYRQYRNSMLSYKALRGLSLSRV ASIFCTTISPDLLLRQCPSHYAIRAGRNLPDKEFRYLRTVITVTAAGWGLSSPLRLR " complement(8715..10223) /product="16S ribosomal RNA" 10753..11397 /locus_tag="CTC00066" 10753..11397 /locus_tag="CTC00066" /codon_start=1 /transl_table=11 /product="conserved protein" /protein_id="AAO34721.1" /db_xref="GI:28202275" /translacion="MKNKKIKLITFIIVIFAYNVKLVNDKINSLPMNDKSIKQMSR KITRDAVVKIQEKQDLITMEVBLNKNKVIDLDSWGNLDKFKKVIISDFYGTGIYTD LAKLNRENVLIQSKTITIRIPKPKVKSITLNEKDTFTKTENGUFRGEEKITPAESQ ILSKAKERNQELNKHVESALSCTKSIEKSIQSIILLPYENINIVIKFIDG" complement(12432..13208) /locus_tag="CTC00067" complement(12432..13208) /locus_tag="CTC00067" /codon_start=1 /transl_table=11 /product="putative polyferredoxin" /protein_id="AAO34722.1" /db_xref="GI:28202276" /translacion="MRGVLCYFGTGNTKMIADKPKYFNMKNIELDIINIIEEKQDME LSEYNFLVVGTPYIGELPKIVDTFLNLPNNGKNIKITILYSTQGTAKASALNVIESKI ENKNYTVSGKIMFKMTNNFYISGKASKEIIEILISCERKVNNTVNNFIEGKYIKE KVNKIRVTLSKIASKHFNHYLNLSKNLSSSENCIRGCLCLRNCPQNNITPENGHVFI HSCKMCLRCIYICPTNEIRYKKNKIDQIKRIIKSLKIK" 13557..14816 /locus_tag="CTC00068" 13557..14816 /locus_tag="CTC00068" /codon_start=1 /transl_table=11 /product="conserved protein" /protein_id="AAO34723.1" /db_xref="GI:28202277" /translacion="MKNLSIYNVIRLLLYFACFIIFTDTPISIFTLKNNIILQPD ILITDYLEISIGIGASFVNSATLCIIVILIRLKLVSIGSVAGLFTVCGFSLFGKQL LNVWPLFLGTWLYSKYQKEPPSKYILVALFGSTLSPAVSQLSFSGHFETSGLIIGIS LGIPLMGFILPISISYCLNHEGYNLNSGPAAGLVGTFPMSILRSFGINFORLWST KYTVLLIIFLPSILFLSMILVFCNLDPSFNKLDKIKHSGKPNKNTIPVWLGTVIAGTINI MGILGIFMWYIILIVDGLNGPIAGTIVTIVGAFGKGFNTIPVWLGTVIAGTINI WSINSPMLLATLFTSTLAPISGVFGPIYGIAGFLHTCIWMNIGHLHAGLNLYNNGF SGGFVALIILPIITNLKVDTONELQI" complement(14871..15398) /locus_tag="CTC00069" complement(14871..15398) /locus_tag="CTC00069"</pre>
gene	<pre>IBEDKIDNIKSTKVNKKVKCDTDEDICTEYKSTEDSKIEDAKVSELHDSILKPD DELVNECEGPGCNFTWTDGRDDDLIK" complement(5249..5324) /product="trNA-Lys" /notes="codon recognized: AAA" complement(5340..5427) /product="5S ribosomal RNA" complement(5606..8450) /product="23S ribosomal RNA" 5938..6594 /locus_tag="CTC00065" 5938..6594 /locus_tag="CTC00065" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAO34720.1" /db_xref="GI:28202274" /translacion="MRRADIEVNPLPVDVDSWGRSACYPRGSFYPLSDGPPTRNHRIT KPDFPCSTCMCRSQAPLCLYLIRAI SDRABGTGRLRYFLGGDRPSPQTAHLTWSRDQ IHGPRLEFOYQGGIPRMPQKLTLLSLPILYRQYRNSMLSYKALRGLSLSRV ASIFCTTISPDLLLRQCPSHYAIRAGRNLPDKEFRYLRTVITVTAAGWGLSSPLRLR " complement(8715..10223) /product="16S ribosomal RNA" 10753..11397 /locus_tag="CTC00066" 10753..11397 /locus_tag="CTC00066" /codon_start=1 /transl_table=11 /product="conserved protein" /protein_id="AAO34721.1" /db_xref="GI:28202275" /translacion="MKNKKIKLITFIIVIFAYNVKLVNDKINSLPMNDKSIKQMSR KITRDAVVKIQEKQDLITMEVBLNKNKVIDLDSWGNLDKFKKVIISDFYGTGIYTD LAKLNRENVLIQSKTITIRIPKPKVKSITLNEKDTFTKTENGUFRGEEKITPAESQ ILSKAKERNQELNKHVESALSCTKSIEKSIQSIILLPYENINIVIKFIDG" complement(12432..13208) /locus_tag="CTC00067" complement(12432..13208) /locus_tag="CTC00067" /codon_start=1 /transl_table=11 /product="putative polyferredoxin" /protein_id="AAO34722.1" /db_xref="GI:28202276" /translacion="MRGVLCYFGTGNTKMIADKPKYFNMKNIELDIINIIEEKQDME LSEYNFLVVGTPYIGELPKIVDTFLNLPNNGKNIKITILYSTQGTAKASALNVIESKI ENKNYTVSGKIMFKMTNNFYISGKASKEIIEILISCERKVNNTVNNFIEGKYIKE KVNKIRVTLSKIASKHFNHYLNLSKNLSSSENCIRGCLCLRNCPQNNITPENGHVFI HSCKMCLRCIYICPTNEIRYKKNKIDQIKRIIKSLKIK" 13557..14816 /locus_tag="CTC00068" 13557..14816 /locus_tag="CTC00068" /codon_start=1 /transl_table=11 /product="conserved protein" /protein_id="AAO34723.1" /db_xref="GI:28202277" /translacion="MKNLSIYNVIRLLLYFACFIIFTDTPISIFTLKNNIILQPD ILITDYLEISIGIGASFVNSATLCIIVILIRLKLVSIGSVAGLFTVCGFSLFGKQL LNVWPLFLGTWLYSKYQKEPPSKYILVALFGSTLSPAVSQLSFSGHFETSGLIIGIS LGIPLMGFILPISISYCLNHEGYNLNSGPAAGLVGTFPMSILRSFGINFORLWST KYTVLLIIFLPSILFLSMILVFCNLDPSFNKLDKIKHSGKPNKNTIPVWLGTVIAGTINI MGILGIFMWYIILIVDGLNGPIAGTIVTIVGAFGKGFNTIPVWLGTVIAGTINI WSINSPMLLATLFTSTLAPISGVFGPIYGIAGFLHTCIWMNIGHLHAGLNLYNNGF SGGFVALIILPIITNLKVDTONELQI" complement(14871..15398) /locus_tag="CTC00069" complement(14871..15398) /locus_tag="CTC00069"</pre>
CDS	<pre>IBEDKIDNIKSTKVNKKVKCDTDEDICTEYKSTEDSKIEDAKVSELHDSILKPD DELVNECEGPGCNFTWTDGRDDDLIK" complement(5249..5324) /product="trNA-Lys" /notes="codon recognized: AAA" complement(5340..5427) /product="5S ribosomal RNA" complement(5606..8450) /product="23S ribosomal RNA" 5938..6594 /locus_tag="CTC00065" 5938..6594 /locus_tag="CTC00065" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAO34720.1" /db_xref="GI:28202274" /translacion="MRRADIEVNPLPVDVDSWGRSACYPRGSFYPLSDGPPTRNHRIT KPDFPCSTCMCRSQAPLCLYLIRAI SDRABGTGRLRYFLGGDRPSPQTAHLTWSRDQ IHGPRLEFOYQGGIPRMPQKLTLLSLPILYRQYRNSMLSYKALRGLSLSRV ASIFCTTISPDLLLRQCPSHYAIRAGRNLPDKEFRYLRTVITVTAAGWGLSSPLRLR " complement(8715..10223) /product="16S ribosomal RNA" 10753..11397 /locus_tag="CTC00066" 10753..11397 /locus_tag="CTC00066" /codon_start=1 /transl_table=11 /product="conserved protein" /protein_id="AAO34721.1" /db_xref="GI:28202275" /translacion="MKNKKIKLITFIIVIFAYNVKLVNDKINSLPMNDKSIKQMSR KITRDAVVKIQEKQDLITMEVBLNKNKVIDLDSWGNLDKFKKVIISDFYGTGIYTD LAKLNRENVLIQSKTITIRIPKPKVKSITLNEKDTFTKTENGUFRGEEKITPAESQ ILSKAKERNQELNKHVESALSCTKSIEKSIQSIILLPYENINIVIKFIDG" complement(12432..13208) /locus_tag="CTC00067" complement(12432..13208) /locus_tag="CTC00067" /codon_start=1 /transl_table=11 /product="putative polyferredoxin" /protein_id="AAO34722.1" /db_xref="GI:28202276" /translacion="MRGVLCYFGTGNTKMIADKPKYFNMKNIELDIINIIEEKQDME LSEYNFLVVGTPYIGELPKIVDTFLNLPNNGKNIKITILYSTQGTAKASALNVIESKI ENKNYTVSGKIMFKMTNNFYISGKASKEIIEILISCERKVNNTVNNFIEGKYIKE KVNKIRVTLSKIASKHFNHYLNLSKNLSSSENCIRGCLCLRNCPQNNITPENGHVFI HSCKMCLRCIYICPTNEIRYKKNKIDQIKRIIKSLKIK" 13557..14816 /locus_tag="CTC00068" 13557..14816 /locus_tag="CTC00068" /codon_start=1 /transl_table=11 /product="conserved protein" /protein_id="AAO34723.1" /db_xref="GI:28202277" /translacion="MKNLSIYNVIRLLLYFACFIIFTDTPISIFTLKNNIILQPD ILITDYLEISIGIGASFVNSATLCIIVILIRLKLVSIGSVAGLFTVCGFSLFGKQL LNVWPLFLGTWLYSKYQKEPPSKYILVALFGSTLSPAVSQLSFSGHFETSGLIIGIS LGIPLMGFILPISISYCLNHEGYNLNSGPAAGLVGTFPMSILRSFGINFORLWST KYTVLLIIFLPSILFLSMILVFCNLDPSFNKLDKIKHSGKPNKNTIPVWLGTVIAGTINI MGILGIFMWYIILIVDGLNGPIAGTIVTIVGAFGKGFNTIPVWLGTVIAGTINI WSINSPMLLATLFTSTLAPISGVFGPIYGIAGFLHTCIWMNIGHLHAGLNLYNNGF SGGFVALIILPIITNLKVDTONELQI" complement(14871..15398) /locus_tag="CTC00069" complement(14871..15398) /locus_tag="CTC00069"</pre>

```
/EC_number="3.6.1.22"
/codon_start=11
/transl_table=11
/product="predicted NADH pyrophosphatase"
/protein_id="AA034724.1"
/db_xref="GI:28202278"
/translation="MKVEYCPKGNKLIKESYWDGQVPCPIDDTMYRDPKPCVVV
AVLKESQILLIKQSYIFENSKVLYSGVTNGETVEETVREVKEEGIEVKNIKYLS
EYUKSEIIMLTFPMABYKSGVIDKSEVEVWNGHIEDALCQMSDEIGKNVVRKVLK
EIDYKGEKAYKCBIK"
/complement(15446..16138)
/locus_tag="CTC00070"
/complement(15446..16138)
/locus_tag="CTC00070"
/codon_start=1
/transl_table=11
/product="sugar fermentation stimulation protein"
/protein_id="AA034725.1"
/db_xref="GI:28202279"
/translation="MIKNKIKRAFIKPNRFOAVVVDNEIMVHPVNTGRCKEIL
IPGTVTLRESNPTKTKYDYLGGYKNGKFINIDSIQPNKVVBEALKGNIKKJXY
NIKREKTYGNRDFPLGNDGSEYLVKGVTLDDGGVTFKFPDAPTERGTKHLEL
IDVKNKIGAGVLFLIQKDVKYFTENIEMDNKFNSEALVSAHKAGVDIFAYECVDVD
FITLKEVKVIL"
/complement(16325..17551)
/locus_tag="CTC00071"
/complement(16325..17551)

gene
CDS

Query Match      23.7%; Score 113.8; DB 1; Length 300029;
Best Local Similarity 56.3%; Pred. No. 3e-19;
Matches 245; Conservative 0; Mismatches 172; Indels 18; Gaps 1;

QY 32 AATTTTATAGAGCATATCTTACAAATCTTACCTCACCGCTATCCCATGCTTTTAGTGG 91
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92 ATAGAATTATAGAGTTACAGCCAAATAAAAAATTGTCGCTTATAGAAATATCACTTTTA 151
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92140 ATAAAGTAGAAAACAGTAGAACCTCGGAAAAAAATTTAGCATATATAAAATGTAACATTTA 92199
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 152 ATCAAGACGTTTAAACGGGCATTTCCCTAATAAGCCATTTTCCCGGGGTTTTCATCG 211
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92200 ATGAAGGATCTTTTAGAGGACATTTCCCGACAGCAAGTAATGCCAGGAGTATTAATTA 92259
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 212 TAGAGGGCATGGCGCAAAACGAGGAGGTTTACCTTCTACCTTCTAGCTTGTGGGGGTTTGACC 271
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92260 TAGAGGATAGCTCAAGCGGGAGCTATAGCAGTATTAAGT-----T 92301
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 272 CTGAAATCGCCAAAACAAAATTCGTGTTATTTTCATGACGATTGATAAGTTAAATCCGCA 331
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92302 TAGACGAATTTAAAGGTAAATATACCATACTTTGCTGGAATAAATAAGCTAAATTTAGAA 92361
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 332 TCCTGTATACCCGAGCGGACAGATTAGAAATACCATTTAGAAAGTCTTAAAGCATAAGGGCA 391
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92362 AGAAAGTTATCCCGGGGACTCTTTAAAAATTAGAGTGGAATGATTAAGTTAAGAGGTT 92421
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 392 TGATCTCGCAAGTGCGTGCACGGCTCAAGTGTGATGCGCAAGTGTGCTGAAGCCGAAT 451
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92422 CAGCAGGTATAGCAGAAAGGTATAGCTAAGTGAATGAAAGGTGGTAGCAGAGCAGAAA 92481
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 452 TGAAGACCATGATTG 466
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92482 TCATGTTTATGATAG 92496
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
AF124757
LOCUS
DEFINITION
ACCESSION AF124757
VERSION
KEYWORDS
SOURCE
Zymomonas mobilis
```

```
ORGANISM      Zymomonas mobilis
REFERENCE     Bacteria; Proteobacteria; Alphaproteobacteria; Spingomonadales;
AUTHORS       Spingomonadaceae; Zymomonas.
TITLE         Lee,H.J. and Kang,H.S.
JOURNAL       Sequence analysis of 43D2 fosmid clone of Zymomonas mobilis ZM4
REFERENCE     Unpublished
AUTHORS       2 (bases 1 to 34083)
TITLE         Lee,H.J. and Kang,H.S.
JOURNAL       Direct Submission
FEATURES      Submitted (29-JAN-1999) Microbiology, Seoul National University,
              San 56-1 Shillindong KwanakGu, Seoul 151-742, Korea
              Location/Qualifiers
              1..34083
                /organism="Zymomonas mobilis"
                /mol_type="genomic DNA"
                /strain="ZM4"
                /db_xref="taxon:542"
                /clone="fosmid 43D2"
                /complement(106..11380)
                /gene="sah"
                /gene="sah"
                /codon_start=1
                /transl_table=11
                /product="succinylarginine dihydrolase"
                /protein_id="AAD29641.1"
                /db_xref="GI:4768847"
                /translation="MIVSEVNFGLIGPTHNVAGLSRGNVASALHAGOPSYPROAALO
                GLEKMKHMDMGLTQGVLPPLRPVTHLHLGVKDDKTILKQAKDRLILPNNILCS
                ASSMWAANAATVISEPFDHGRVHPTANTLATMLHRHLEAQTTYAQLNQMFNSCFPA
                MHPLPCQHFSDEGAANMRITSAHGRGTGNIIVFYDKNDIYPARQKLRSQAVALR
                GEVYKDLAFIPQKEAIAKAGFNDVAVANETVYLLAHAEAFEDQGWIKRIAEKID
                GFPIIDNTLEQAVKSYLFNSQIVTLKDRTMALILPQEVKSDPAVWETVNRILISGN
                NPQIAVVVDVRESMWANGGPACLRVPLSLNALEAVDQRFILTPKRWEMLYQLVEN
                FWDKTPDLDLVDLMLKTAVRHAWALTSWLG"
              1506..2540
                /note="hypothetical protein; ORF1"
                /codon_start=1
                /transl_table=11
                /product="unknown"
                /protein_id="AAD29642.1"
                /db_xref="GI:4768848"
                /translation="MKPEVFIQIADRLSDQEKLFROYEPAGYILFORNITDRTQLKK
                LTADLANSQONTIPILIDQEGGRVARKMPPVMPAPFAAARFDELYQKAPVSAIEAVRA
                NSHALALLQDVGINDVCDMLDIRDPEDGNIIGDRSGHNPEOVAAALGRAVLGLAK
                GSVGCIKHIPGHQAKVDSHKALPVDSPIELERDIRPQHLENANAKMTAHLIYQ
                AWKSCASYSYKIIHDIIRKKGILGMLSDDIGMNALDGPVTRALRVFDAGCDVA
                LHCSGFPDDMKAILENNDEIRPESFAALEKVTFNEFSASDFDLDRANKALPDRCIAE
                RDALFLEAAV"
                /complement(2537..2845)
                /note="hypothetical protein; ORF2"
                /codon_start=1
                /transl_table=11
                /product="unknown"
                /protein_id="AAD29643.1"
                /db_xref="GI:4768849"
                /translation="MNTAFDITTLTAKADQEQALQOALLHYADQSKTEGGLAFH
                IGDKSQPPPIILNERWQDHAAPFAHEKSDIFQDFASMFKRFLEKPLLSVSVPLMED"
              3954..5909
                /gene="sah"
                /gene="sah"
                /gene="sah"
                /codon_start=1
                /transl_table=11
                /product="glutaryl-7-ACA acylase precursor"
                /protein_id="AAD29644.1"
                /db_xref="GI:4768850"
                /translation="MSRLKLSVNLISAMKKYLMRGSVVASLTSILAIPALSNAPDKTIW
                PENGDIAMFSAPTAHYDEKREVMIPMRGVKHLTVIVPKNRNLFILITRTPYDA
                SSRTSRSDPSMLATLPEGDEYFVRDGYIRVQDVRGKYGSEGVYVTRPVFGDNLNPT
                KVDHTTDAWDTLEWLTKHIPESGNRVGMIGSSYEGFTVAALLNPHPPALKVAAAPSPM
```

```
VDGMGDDMFHYGAFRAAFDYFLROMTAKGTGAPPVHGAYDDYKAFLETSGAGNNAK
KEGIDQIPWQRLSIHPAYDKWQOQALDQIVAAHPSPVPTWVIGLWDQEDMWGAIH
SFEGLKQAGHIDTNYLVGPNWRHSQVNTYNGSLGALHWDGDTALQFRDRDTLLPFNNRY
SFKLQAPAEYNTYNGENHDKLNDWQDKKQLTPLYLOAQASFTKPTSGNA
PSDOYISDPKPKVPYLPITFADTRWKMILIEDORFAASRPDVLTYETPLVDHPEK
LRGAPPANLLAATGSDVDWVKLIDVYDEIPSDPKMGQQLAISMDIFRGIRYNSF
EKSPSPGAKGVQQRFRFLPVVDHVPFLPGRHIRMVQIQSSLFPLIDRNRPORYVENIFAK
PADYAAVTETVMHSPDQASSVELFVP"
6171..7190
/gene="reca"
CDS
6171..7190
/function="recombinase"
/codon_start=1
/transl_table=1
/product="ReCA protein"
/protein_id="AAD29645.1"
/db_xref="GI:4768851"
/translation="MDKQKALEAALQAIDRAFKGSVMLGSRGKEIEDTITSSGSLGL
DIALIGGLPRGRIVEIYSPSSGKTLIALHAIABQAKAGTAAFDVAEHALDPYAK
KLGVNTDOLIVSDPTGQALIEITDLIRSNVADILVDSVAALTPRAEIEGMDGSH
VGLQARLMSQALRKITGINSQTLVFINOVRMKIGVMYGNPETTTGNALKFVASV
RLDIRVQIQIDRDEIVGNATRVKVKLLRPFKQVEFIDMVGSGVSKWGLDILGV
KAGIDKSGSWFSHDSVRIGORENAKTLRHPENTEKIEKMRHRTAIEADEMLDP
SIPED"
7338..8114
/gene="comp"
CDS
7338..8114
/gene="comp"
/codon_start=1
/transl_table=1
/product="competence protein F"
/protein_id="AAD29646.1"
/db_xref="GI:4768852"
/translation="MPYHAAQIVARSFWLHSPFCVCSWOSLHFLVPACIOCGPPLP
RDOQTEPCGLOHPSPFRDMAAAYDEISRQLIKCKYQKQMLAKDMAKLMWRFL
PFIKSNIRAVQNQLADKCPNPPSSSPQTPLEIIPVLRWRLWQRFNGFNQAALIQY
LAKAGSYDDCDLSLYRRHSKPLGLHTPKRQHWKNVAFGLRIGKQRDLGRDIIIFI
DDVFTSGATTSCARLKKAGVSVHVLWCARVLPQDNFL"
complement(8200..8895)
/gene="phoB"
CDS
complement(8200..8895)
/gene="phoB"
/codon_start=1
/transl_table=1
/product="phosphate regulatory protein PhoB"
/protein_id="AAD29647.1"
/db_xref="GI:4768853"
/translation="MAALRLILVEDDDALAEILLKWHFSKQNFVHVHTRDGEALLLAR
ENPPDMLDMWVEGTSGIEVCQRQLRRYPETANLPIIMLTARGEEDDRIRGLETGADD
YVTKPSPRELVAVLVLRIRPALSKEQLNVDIEMDLVSHRVKRACTNYSIGPTE
FRLLRHLMYPRVRSREKLLDIYIWGQSDIDELRTVDVHIRLRKALNRDGLPDLIRT
VRSAGYALDTDEE"
complement(8895..9608)
/gene="phoU"
CDS
complement(8895..9608)
/gene="phoU"
/codon_start=1
/transl_table=1
/product="transcriptional regulator PhoU"
/protein_id="AAD29648.1"
/db_xref="GI:4768854"
/translation="MKKASEIGHYKAFDEDLGLRALICEMGGLTEIALTSTLEALI
HSNKELAAQIVSKQKEIDALEERARVAMRIISLRAPLADDLRDVIAANKISGILERM
ADYAKNIARVPILQNMVSIHPVPLVPVNGEMVSMVRNVLDAYAADPKKALHVAEY
DMEVDGLYNLFRILLTYMMEDSGRISTCIHLMFIKRLNLERIGDHATNIAEMVYFAAT
GKQMPERTRGKDLLASEGS"
complement(9815..11080)
/gene="phoR"
CDS
complement(9815..11080)
/gene="phoR"
/codon_start=1
/transl_table=1
```

```
/product="phosphate regulon sensor protein PhoR"
/protein_id="AAD29649.1"
/db_xref="GI:4768855"
/translation="MTFPQSRKPKDCVIRKIPVYLSGLSILALFIIAVSHSVGWPIL
SLMLIVILGMWFKNISEATPPKSKSGNSLTDPAFPQOATAADIIIAMDEPALLV
NOGOVEVANRAENALLGMHIEGGDIRMAIRHPAAIRLLTRPDLNMTVLLAGLGQVNR
RVELIYAPLNDEQRLLILLRDOQSTAHILTEQVRIDPFAVNTSEHLRPLATLSGFIELED
VDKIKERDTHHFLSIMAREAKQMLNDLMSLSRTEAGKFSPLHDIIVHMSPVVEEV
VNIQASGQAKATQITLENHSSDSLQSDRAELTQLLYNIIGNALKYRKDGIKVAL
DNTDDHRIKLAIADEGDGIPYHHIPRLTERFYRVDSRSRALGCTGLGLAIVKQIVEH
HRGELLDSLPGKGTVTTVFLPTTDFPE"
11284..12006
/notes="ORF3"
CDS
11284..12006
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAD29650.1"
/db_xref="GI:4768856"
/translation="MNKEQRKNSAQSIRWHPSPHPEGRKFFVIAVIIISALITYFS
WLMWAMPLGVLCFAAFPRDPIRTVPEGGLIVSPADGMVCLLADVPPPELAGADG
LGDAPLTRYSIFMSVDFHINRAPVRIARIAYIPGAFVNADLDKASEKNERQHFI
ETKEGINVGLTQAGLIARRIVPLTKVEDYVERGERIGLIRFGSRLDVLYPAGVTPQV
ALGQRCLAGETFWDRLDHKPSR"
12147..12821
/gene="psaA"
CDS
12147..12821
/gene="psaA"
/codon_start=1
/transl_table=1
/product="CDP-diacylglycerol--serine
O-phosphatidyltransferase"
/protein_id="AAD29651.1"
/db_xref="GI:4768857"
/translation="MRYPVEWEHAIYAILVAALVDLDRVARLLKGATRGAEILDSL
SDAIEQVAPAILVYLSLQNWPRWGLFALSHAVCCALRLARFNVNIDYLDOPKSA
GFITGPAPAGNFSWMPLYANTVSYVDLFRMPELLVGLWLSLAFMLADIPLTEGMSV
LRVPRKLUGLLVYVAAFAAISAPWAVLSVFSVIYTMFLDPFGILSYKSMKSKEKT
IVAEES"
complement(12976..13674)
/gene="asp"
CDS
complement(12976..13674)
/gene="asp"
/codon_start=1
/transl_table=1
/product="aspartate racemase"
/protein_id="AAD29652.1"
/db_xref="GI:4768858"
/translation="MKKGLIGIGPESTLLYYKKLYVGANNRIGSQFPNLTVESLN
```

```
Query Match 22.4%; Score 107.6; DB 1; Length 34083;
Best Local Similarity 55.7%; Pred. No. 1.4e-17;
Matches 234; Conservative 0; Mismatches 174; Indels 12; Gaps 1;

Qy 60 TCTACCTCAGCGGTATCCCATGCTTTTAGTGATAGATAATATTAGAGTTACAGCCATAA 119
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25278 TCTGCCTCAGCGTTTCCGATGCTGCTGTTGATCGGGTCGAGACACTGATTCGGGCGA 25337

Qy 120 AAAAAATTGCTTATAAGATATACATTTTAAAGAACGCTGTTTAAACGGGCATTTCCC 179
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25338 AAAAAATTGTCGATCAAGCGCGTAACAAATTAACGAGCGGTCCTTTACGGTCATTTCCC 25397

Qy 180 TAAATAAGCCCATTTTCCCGGCGCTTTTTCATCGTAGAGGCATGCGCAAAACGGAGGGTT 239
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25398 CGGTAAATCCGATATCCAGGTGTTTTAAATGTGCGAGCGCTGGCACAAGCTGCCGGTGT 25457

Qy 240 TTTAGCCTTCACATGCTGTGGGGGTTTGACCCCTGAATCGCCAAAAACAAAATCGTGTA 299
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25458 TCTTGCCATTGAAAGCTTAGGACTG-----ACCGGAACCGGTAAAGCTGTTTA 25505

Qy 300 TTTTCATGACGATTGATAAGGTTAAATTTCCGCATCCCTGTAAACCCAGCGGACAGATTAGA 359
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25506 TTTTCATGCGCTATTAAACGAGACAAAATTCGCTATTCCCGGTAGAACCGGGTATTTTATTGCG 25565
```

QY 360 ATACCATTTAGAACTCTTAAGCATTAAGGCGCATGCTGCGAAGTGGTGGCAAGGCTCA 419
Db 25566 GTTAGAAGTCGAGTTTTCACAAAACGGGCAAGAATTTGAATTTAAAGCCCGTGCCTT 25625
QY 420 AGTGGGATGCAAAAGTGGTCTGAAAGCCGAATTTGAAAGCCATGATTCGACAGAGAGATTA 479
Db 25626 GATTGAAGACAAAGTCGGCTGAACAGAGATTCACCGCCATGATCGCTGATCCAGCTAA 25685

RESULT 12

AE008692.11/c

WPCOMMENT

Sequence split into 21 fragments LOCUS AE008692 Accession AE008692

Fragment Name	Begin	End
AE008692_00	1	110000
AE008692_01	100001	210000
AE008692_02	200001	310000
AE008692_03	300001	410000
AE008692_04	400001	510000
AE008692_05	500001	610000
AE008692_06	600001	710000
AE008692_07	700001	810000
AE008692_08	800001	910000
AE008692_09	900001	1010000
AE008692_10	1000001	1110000
AE008692_11	1100001	1210000
AE008692_12	1200001	1310000
AE008692_13	1300001	1410000
AE008692_14	1400001	1510000
AE008692_15	1500001	1610000
AE008692_16	1600001	1710000
AE008692_17	1700001	1810000
AE008692_18	1800001	1910000
AE008692_19	1900001	2010000
AE008692_20	2000001	2056415

Continuation (12 of 21) of AE008692 from base 1100001 (AE008692 Zymomonas mobilis subsp.

Query Match 22.4%; Score 107.6; DB 1; Length 110000;
Best Local Similarity 55.7%; Pred. No. 1.4e-17;
Matches 234; Conservative 0; Mismatches 174; Indels 12; Gaps 1;
QY 60 TCTACTCACCCTATCCCATGCTTTTATGAGTAAATATAGATTACAAGCCATATA 119
Db 63029 TCTGCCCTCACCGCTTTTCCGATGCTGTGATCGGTCGAGACACTGATTCGGGCGCA 62970
QY 120 AAAAATTGTCGCTTATAAGAAATATCACTTTTAATGAAGACGCTTTAAACGGGCAATTCCC 179
Db 62969 AAAAATTGTCGATCAAGCGCGTAAACAATTACGAGCGCTTCTTTACCGGTCATTTCCC 62910
QY 180 TAATAAGCCCATTTTCCCGGGCGTTTGTATCGTAGAGGCGATGGCGCAAAACGGGAGGGTT 239
Db 62909 CGGTAATCCGATTATGCCAGGTGTTTAATTTGTCGAGGCTTGGCACAGCTGCCGGTGT 62850
QY 240 TTTAGCCTTCACTAGCTTGTGGGGGTTTGACCTGAAATTCGCAAAACAAAATTCGTGTA 299
Db 62849 TCTTGCCATTGAAAGCTTAGGACTG-----ACCGGAACCGGTAAGCTGGTTA 62802
QY 300 TTTTCATGAGATTGATAAGTTAAATTCGCATCCCTGTAAACCCAGGCGACAGATTAGA 359
Db 62801 TTTTCATGGCTATTAAACGACACAAAATTCGCTATTCCGGTAGAACCGGGTATTTTATTGG 62742
QY 360 ATACCATTTAGAACTCTTAAGCATTAAGGCGCATGCTGCGAAGTGGTGGCAAGGCTCA 419
Db 62741 GTTAGAAGTCGAGTTTTCACAAAACGGGCAAGAATTTGAATTTAAAGCCCGTGCCTT 62682
QY 420 AGTGGATGCAAAAGTGGTCTGAAAGCCGAATTTGAAGCCATGATTCGACAGAGAGATTA 479
Db 62681 GATTGAAGACAAAGTCGGCTGAACAGAGTTTCCCGCCATGATCGCTGATCCAGCTAA 62622

RESULT 13

AE017180.24/c

WPCOMMENT

Sequence split into 39 fragments LOCUS AE017180 Accession AE017180

Fragment Name	Begin	End
AE017180_00	1	110000
AE017180_01	100001	210000
AE017180_02	200001	310000
AE017180_03	300001	410000
AE017180_04	400001	510000
AE017180_05	500001	610000
AE017180_06	600001	710000
AE017180_07	700001	810000
AE017180_08	800001	910000
AE017180_09	900001	1010000
AE017180_10	1000001	1110000
AE017180_11	1100001	1210000
AE017180_12	1200001	1310000
AE017180_13	1300001	1410000
AE017180_14	1400001	1510000
AE017180_15	1500001	1610000
AE017180_16	1600001	1710000
AE017180_17	1700001	1810000
AE017180_18	1800001	1910000
AE017180_19	1900001	2010000
AE017180_20	2000001	2110000
AE017180_21	2100001	2210000
AE017180_22	2200001	2310000
AE017180_23	2300001	2410000
AE017180_24	2400001	2510000
AE017180_25	2500001	2610000
AE017180_26	2600001	2710000
AE017180_27	2700001	2810000
AE017180_28	2800001	2910000
AE017180_29	2900001	3010000
AE017180_30	3000001	3110000
AE017180_31	3100001	3210000
AE017180_32	3200001	3310000
AE017180_33	3300001	3410000
AE017180_34	3400001	3510000
AE017180_35	3500001	3610000
AE017180_36	3600001	3710000
AE017180_37	3700001	3810000
AE017180_38	3800001	3814139

Continuation (25 of 39) of AE017180 from base 2400001 (AE017180 Geobacter sulfurreducens

Query Match 22.2%; Score 106.4; DB 1; Length 110000;
Best Local Similarity 53.9%; Pred. No. 2.8e-17;
Matches 249; Conservative 0; Mismatches 201; Indels 12; Gaps 1;

QY 19 AACTTGCAATCTCAATTTTATAGAGCATATCTTACAAATTTCTACCTCACCGCTATCCC 78
Db 81306 AACATGGAACAGTCTTTTGATATCAACGAATCATGAAGATTTCTCCCGCACCGCTACCCC 81247
QY 79 ATGCTTTTAGTGATAGATATATAGATTACAGCCATTAATAAATAATGTCGCTTATAG 138
Db 81246 TTCTCTCTGTGGATCGGATTTGGAGTACGTTGCCGCGAGCGGATCGTGGGGATCAAG 81187
QY 139 AATATCACTTTTAATGAAGACGCTGTTAAACGGGCATTTCCCTAATAAGCCCATTTTCCCG 198
Db 81186 AACGTGAGCATCATGAGCCGTTTTTCCAGGCCACATTTCCGGGACACCCCGGTGATGCG 81127
QY 199 GCGTTTTTGTATCGTAGAGGCGATCGCGCAACACGGAGGGTTTTTATAGCTTTCATAGCTTG 258
Db 81126 GGAGTCTGTATCGTTGAGGCCATGCGCCAGTGGGGGAATCTACGCCCTATGTGACCTTG 81067
QY 259 TGGGGGTTTGAACCTGAATTCGCCAAACAAAATCGTGATTTTCATGACCATGATAG 318
Db 81066 GCGGACGAGGTCCGCGA-----CAAGGCTCTGCTACTTTTGGCTTCCATCGACAAC 81019
QY 319 GTTAAATTCGCGATCCCTGTAAACCCAGGCGACAGATTTAGAAATACCATTTTGAAGTCTTA 378
Db 81018 GTCAAGTTCCGCAAAACCGGTGGTCCCGGAGACCAACTCGCGGATCGAAGTACGATCTCC 80959
QY 379 AAGCATAAGGGCATGATCTCGCAAGTGGGTGGCACCGCTCAAGTGGATGGCAAGTGGTC 438

FEATURES
source

Location/Qualifiers
1. 329100
/organism="Sinorhizobium meliloti"
/mol_type="genomic DNA"
/strain="1021"
/db_xref="taxon:382"
152. .227
/gene="tRNA-THR_TGT"
132. .227
/gene="tRNA-THR_TGT"
/product="tRNA-Thr"
/note="codon recognized: ACA; predicted by tRNAscan-SE"
/evidence=not experimental
complement(373. .798)
/gene="SMC04435"
complement(373. .798)
/gene="SMC04435"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Homology"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAC45913.1"
/db_xref="GI:15074267"
/db_xref="InterPro:IPR004360"
/db_xref="InterPro:IPR011588"
/db_xref="UniProt/TREMBL:Q92Q19"
/translation="MNNALVPEFVSDWQAKSEFYCNILGFCSDYERPEEGFCYLSLD
GAELMDIGTGRFTADGHLPSYFPGRLNQVIRVASADALVRALTDRGIALYPLF
ERWYRGDEYGVNPDVDPDGYLLRFYELGRPRPTS"
complement(914. .1855)
/gene="SMC01330"
complement(914. .1855)
/gene="SMC01330"
/function="miscellaneous; not classified regulator"
/note="Product confidence : putative
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="PUTATIVE TRANSCRIPTION REGULATOR PROTEIN"
/protein_id="CAC45914.1"
/db_xref="GI:15074268"
/db_xref="GOA:Q92Q18"
/db_xref="InterPro:IPR000847"
/db_xref="InterPro:IPR005119"
/db_xref="InterPro:IPR011991"
/db_xref="UniProt/TREMBL:Q92Q18"
/translation="MQDLNDLFAFVVKHNGFSAARDLNVPKSLKHVARLEBQL
GVRLRSTRKLWTEGRI FYEHAQGLDGVAAAEIAA VRAEPTGVVRLACPIGF
TPMLADILPAFRYPGVRLIATNSRIDIIEERIDVALRARDOLDTDSQIIVKFG
VRODLAASPTLLARGLITDNLDSQMTLSNNEQHPNDVRLVHATGAEIAHAPV
EVCSDFLLEAALAEAGMGIALLPDHCIRAFRTGALVPLPWTSGNVVHVLVFPSSR
GLLPATRALIDFLAENLIKALERCREDVPRPAASFEI"
2063. .2683
/gene="acpD"
/note="synonym: SMC01329"
2063. .2683
/gene="acpD"
/function="small molecule metabolism; fatty acid
biosynthesis"
/note="Product confidence : probable
Gene name confidence : putative
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1

/evidence=not experimental
/transl_table=11
/product="PROBABLE ACYL CARRIER PROTEIN PHOSPHODIESTERASE"
/protein_id="CAC45915.1"
/db_xref="GI:15074269"
/db_xref="GOA:Q92Q17"
/db_xref="InterPro:IPR003680"
/db_xref="UniProt/Swiss-Prot:Q92Q17"
/translation="MKILHIDSGILGHSVSRRLTSAIVSQLKADRPDAEITYRDLAS
ERVPHLTGAQIMAPADLEGVDALLAADVRIGRQMLLEFLAADTVVVGAPMNFSPSQ
LKAVIDLAVAGKTFRYTEAGPEGLARKKLIIVASTRGHYSVAPASAMDHQETYLRS
VFQFFGTDIEFTRAEGNLNLPDQKQFALIAEAKTIAEGNVKLAS"
2792. .3076
/gene="SMC01328"
2792. .3076
/gene="SMC01328"
/function="miscellaneous; hypothetical/partial homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Framed"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="HYPOTHETICAL TRANSMEMBRANE PROTEIN"
/protein_id="CAC45916.1"
/db_xref="GI:15074270"
/db_xref="GOA:Q92KE9"
/db_xref="UniProt/TREMBL:Q92KE9"
/translation="MKCVRLRFLALSAAVALVAGSVASAPGRCCTCRHDKAYE
LQQTACIRIGDISYLARCENNLVNTWKRLQDGCPTAGLAVTVVPSAIP"
complement(3066. .3941)
/gene="SMC01327"
complement(3066. .3941)
/gene="SMC01327"
/EC_number="2.1.1.1"
/function="macromolecule metabolism; macromolecule
synthesis, modification; rna synthesis, modification , dna
transcription"
/note="Product confidence : putative
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="PUTATIVE TRNA/RNA METHYLTRANSFERASE PROTEIN"
/protein_id="CAC45917.1"
/db_xref="GI:15074271"
/db_xref="GOA:Q92Q16"
/db_xref="InterPro:IPR001537"
/db_xref="InterPro:IPR004441"
/db_xref="UniProt/TREMBL:Q92Q16"
/translation="MSKDKTGKSARDYHNNRRAHRARRRERGEIPTPRQDKRAKL
SADWKPALAPQVLLYGLHVTAAADNPERRIIRLSVTQNAARLDGLTALPFPPT
ETVMPQLDRILGPDAIHOGWMLSTQPLPHRKEALRDCPLILVLDVTPHNVGAIM
RAVAFDAAGALITTMHRSPTESGLVSLAKSLEALDIPYIQTINLADTLBELHGLFMT
IGLDSGPOPLEGTFFAGDKIALVLGSEGGKLRQKTRQTVKALARLDMPGAIKSLNVS
AAAIAMYAAARHLKG"
4158. .4242
/gene="tRNA-TYR_GTA"
4158. .4242
/gene="tRNA-TYR_GTA"
/product="tRNA-Tyr"
/note="codon recognized: UAC; predicted by tRNAscan-SE"
/evidence=not experimental
4266. .4339
/gene="tRNA-GLY_TCC"
4266. .4339
/gene="tRNA-GLY_TCC"
/product="tRNA-Gly"
/note="codon recognized: GGA; predicted by tRNAscan-SE"


```
gene /evidence=not_experimental
4602..5777
/gene="tufB"
/note="synonym: SMC01326"
CDS 4602..5777
/gene="tufB"
/function="macromolecule metabolism; macromolecule
synthesis, modification; proteins and peptides -
translation and modification"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not_experimental
/transl_table=11

Query Match 21.6%; Score 103.8; DB 1; Length 329100;
Best Local Similarity 56.7%; Pred. No. 1.3e-16;
Matches 242; Conservative 0; Mismatches 167; Indels 18; Gaps 2;

QY 40 ATAGAGCATATCTTACAAATTTCTACCTCAGCGTATCCCATGCTTTTGTGGATAGAAAT 99
Db 165453 ATACAGGAAATCCTGAGGCTTCTTCCCATCGTTACCCCTTCCTGCTGGTGGACCGCATC 165512

QY 100 ATAGAGTTACAAGCCAATAAAAAATTTGTCGCTTATAAGAAATATCACTTTTAATGAAGAC 159
Db 165513 ATCGAGATCGATGACGACAAATTCGGCCATCGGGATCAAGAACGTGACGGCCACGAGCCG 165572

QY 160 GTGTTTAAAGGGCATTTCCCTAATAAGCCCAATTTCCCGGGCGTTTGTATCGTAGAGGGC 219
Db 165573 CATTTTCAACGGGCACATTTCCCGAGAGCCGATCATCGCGGCGTTCTTCTGATCGAAGGC 165632

QY 220 ATGGCGCAACCGGAGGGTTTTTAGCCTTCACTAGCTTGTGGGGTTTGACCTGAAATC 279
Db 165633 ATGGCGCAGACGGCCGG-----CGCGATCTGTGCGGCAAGACGGGAATC 165677

QY 280 GCCAAACAAAATCGTGATTTTCATGACGATTGATAAGTTAAATTCGCGATCCCTGTA 339
Db 165678 G---GAAGCAATCTCGTCTATTTTCATGACGATCGACAATGACGGTTCCGCAACCCGTC 165734

QY 340 ACCCCAGGCGACAGATTAGAAATACCAATTTAGAACTCTTAAAGCATAAAGGGCATGATCTGG 399
Db 165735 GTGCCGGGCGACCGGTGGAGTTCATGTCGTGAAGCAGAGCAGCGCGGCAATATCTGG 165794

QY 400 CAAAGTGGTGGCAGCGCTCAAGTGGATGGCAAGTGGTCTGTAAGCCGAATTTGAAGCC 459
Db 165795 AAATTTCACTGGCATGCAAAAGTTGACGGCAACTGTCGCGGAAGCTGATATCGGCGCG 165854

QY 460 ATGATTG 466
Db 165855 ATGATTG 165861
```

Search completed: January 12, 2006, 02:20:52
Job time : 3018 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 11, 2006, 21:53:19 ; Search time 461 Seconds
(without alignments)
6939.376 Million cell updates/sec

Title: US-10-662-126-36

Perfect score: 480

Sequence: 1 atggacaaagccatcaaaaa.....tgattcgagagagattaa 480

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 496997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	480	100.0	480	8	ACA34928
2	454.4	94.7	480	4	AAS53867
3	294	61.3	853	6	ABX66131
4	134.4	28.0	441	8	ACA30215
5	100	20.8	659	6	ABX66187
6	99.2	20.7	459	13	ADA59846
7	95	19.8	432	8	ACA28687
8	93.4	19.5	417	8	ACA33871
9	92.6	19.3	462	8	ACA53362
10	88.6	18.5	435	8	ACA36484
11	88.6	18.5	1755	6	ABQ69109
12	88.6	18.5	1927	6	ABQ70710
13	88.6	18.5	110000	6	ABA03041_25
14	88.6	18.5	110000	6	ABA03041_26
15	88.2	18.4	459	8	ACA34312
16	87.4	18.2	453	8	ACA26987
17	87	18.1	110000	6	ABQ67196_5
18	87	18.1	110000	6	ABQ69245_26
19	86.6	18.0	453	10	ACF69720

C	20	86.6	18.0	110000	10	ACF67367_27	Continuation (28 o
	21	86.6	18.0	110000	10	ACF65386_1	Continuation (2 of
	22	86.4	18.0	426	4	AAS53189	Aas53189 Enterococ
	23	86.4	18.0	429	10	ADH83263	Adh83263 Enterococ
	24	86.4	18.0	6021	2	AAX13395	Aax13395 Enterococ
	25	86.4	18.0	6021	6	ABS99190	Abs99190 Enterococ
	26	86.4	18.0	243335	14	ABE42735	AbE42735 L. pneumo
	27	86.4	18.0	298667	14	ABE39173	ABE39173 L. pneumo
	28	86.2	18.0	414	13	ADS63118	AdS63118 Bacterial
	29	86.2	18.0	414	13	ADS62218	AdS62218 Bacterial
	30	86.2	18.0	414	13	ADS62656	AdS62656 Bacterial
	31	86.2	18.0	414	13	ADS59716	AdS59716 Bacterial
	32	86.2	18.0	426	8	ACA27788	ACA27788 Prokaryot
	33	85	17.7	423	8	ACA33329	ACA33329 Prokaryot
	34	85	17.7	471	13	ADS45464	AdS45464 Bacterial
	35	85	17.7	518	13	ADT46454	AdT46454 Bacterial
	36	84.8	17.7	110000	14	ABE39175_06	Continuation (7 of
	37	84.8	17.7	295644	14	ABE35721	AbE35721 L. pneumo
	38	83.6	17.4	420	13	ADT45905	AdT45905 Bacterial
	39	83.2	17.3	110000	14	ABE42401_05	Continuation (6 of
	40	82.8	17.2	438	8	ACA45733	ACA45733 Prokaryot
	41	82.8	17.2	441	4	AAS54202	Aas54202 Pseudomon
	42	82.8	17.2	441	8	ACA42462	ACA42462 Prokaryot
	43	82.8	17.2	441	10	ADE49810	Ade49810 (3R)-hydr
	44	82.8	17.2	441	10	ADE49808	Ade49808 (3R)-hydr
C	45	82.8	17.2	465	11	ABD05628	ABD05628 Pseudomon

ALIGNMENTS

RESULT 1	ACA34928
ID	ACA34928 standard; DNA; 480 BP.
XX	
AC	ACA34928;
XX	
DT	19-JUN-2003 (first entry)
XX	
DE	Prokaryotic essential gene #16585.
XX	
KW	Antisense; ds; prokaryotic essential gene; cell proliferation;
KW	drug design; gene.
XX	
OS	Helicobacter pylori.
XX	
PN	WO20027183-A2.
XX	
PD	03-OCT-2002.
XX	
PF	21-MAR-2002; 2002WO-US009107.
XX	
PR	21-MAR-2001; 2001US-00815242.
PR	06-SEP-2001; 2001US-00948993.
PR	25-OCT-2001; 2001US-0342923P.
PR	08-FEB-2002; 2001US-00072851.
PR	06-MAR-2002; 2002US-0362699P.
PA	(ELIT-) ELITRA PHARM INC.
XX	
PI	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX	WPI; 2003-029926/02.
DR	P-PSDB; ABU31058.
XX	
PT	New antisense nucleic acids, useful for identifying proteins or screening
PT	for homologous nucleic acids required for cellular proliferation to
PT	isolate candidate molecules for rational drug discovery programs.
XX	
PS	Claim 14; SEQ ID NO 22798; 1766pp; English.
XX	
CC	The invention relates to an isolated nucleic acid comprising any one of

the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 480 BP; 150 A; 93 C; 107 G; 130 T; 0 U; 0 Other;

Query Match 100.0%; Score 480; DB 8; Length 480;
Best Local Similarity 100.0%; Pred. No. 1.1e-137;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGAAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
1 ATGGAAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
61 CTACCTCACCCTATCCCATGCTTTTATAGAGTATAGATTAAGCAAGCAATAA 120
61 CTACCTCACCCTATCCCATGCTTTTATAGAGTATAGATTAAGCAAGCAATAA 120
121 AAAATTGCTGCTTATAGATATACATCTTTTATAGAGAGTGTTTAACGGGCATTTCCCT 180
121 AAAATTGCTGCTTATAGATATACATCTTTTATAGAGAGTGTTTAACGGGCATTTCCCT 180
181 AATAAGCCATTTTCCGGCGCTTTTATAGAGAGGATGCGCAAAACGGGAGGTTT 240
181 AATAAGCCATTTTCCGGCGCTTTTATAGAGAGGATGCGCAAAACGGGAGGTTT 240
241 TTAGCCTTCACTAGCTTTGGGGGTTTGCACCTGAAATGCCAAAACAAAATCGTGAT 300
241 TTAGCCTTCACTAGCTTTGGGGGTTTGCACCTGAAATGCCAAAACAAAATCGTGAT 300
301 TTATGAGCATTTGATAGGTTAAATTCGGCATCCCTGTAAACCCAGCGCAGATTAGAA 360
301 TTATGAGCATTTGATAGGTTAAATTCGGCATCCCTGTAAACCCAGCGCAGATTAGAA 360
361 TACCATTTAGAGCTTTTAAGCATATAGGCGATCATCTGGCAAGTGGTGCAGGCTCAA 420
361 TACCATTTAGAGCTTTTAAGCATATAGGCGATCATCTGGCAAGTGGTGCAGGCTCAA 420
421 GTGGATGGCAAGTGGTGGTGAAGCGCAATTCGAAAGCCATGATTCGAGAGAGATTAA 480
421 GTGGATGGCAAGTGGTGGTGAAGCGCAATTCGAAAGCCATGATTCGAGAGAGATTAA 480

RESULT 2
AAS53867
IP AAS53867 standard; DNA; 480 BP.

AAS53867;
13-FEB-2002 (first entry)
Helicobacter pylori DNA for cellular proliferation protein #321.
Antisense; ds; prokaryotic cellular proliferation gene; antibiotic; antibacterial; drug design.
Helicobacter pylori.
WO200170955-A2.
27-SEP-2001.
21-MAR-2001; 2001WO-US009180.
21-MAR-2000; 2000US-0191078P.
23-MAY-2000; 2000US-0206848P.
26-MAY-2000; 2000US-0207727P.
23-OCT-2000; 2000US-0242578P.
27-NOV-2000; 2000US-0253625P.
22-DEC-2000; 2000US-0257931P.
16-FEB-2001; 2001US-0269308P.
(ELIT-) ELITRA PHARM INC.
Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
Yamamoto RT, Xu HH;
WPI; 2001-611495/70.
P-PSDB; AAU36008.
New polynucleotides for the identification and development of antibiotics, comprise sequences of antisense nucleic acids.
Claim 27; SEQ ID NO 7504; 511pp; English.
The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes, themselves and the encoded proteins. The prokaryotes used are *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence encodes an essential prokaryotic cellular proliferation protein. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
Sequence 480 BP; 151 A; 94 C; 107 G; 128 T; 0 U; 0 Other;
Query Match 94.7%; Score 454.4; DB 4; Length 480;
Best Local Similarity 96.7%; Pred. No. 9e-130;
Matches 464; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
Db 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAGATT 60
Qy 61 CTACCTCACCCTATCCCATGCTTTTATAGTGAAGAATTATAGAGTACAGCCCAATAA 120
Db 61 TTACCTCACCCTATCCCATGCTCTTTAGTGAAGATTATAGAGTACAGCCCAATAA 120
Qy 121 AAAATTGCTGCTTATAGATATACATCTTTTATAGAGAGTGTTTAACGGGCATTTCCCT 180

Db 121 AAAATTGCGCTTATAGATAATACACTTTTAAATGAAGATGTTTAAACGGGCATTTCCCT 180
QY 181 AATAAGCCCATTTTCCGGGCGTTTTCATCGTAGAGGGCATGCGCAACCGGAGGGTTT 240
Db 181 AATAAGCCCATTTTCCGGGCGTTTTCATCGTAGAGGGCATGCGCAACCGGAGGGTTT 240
QY 241 TTAGCCCTTCACTAGCTTGTGGGGGTTTGACCCCTGAATCGCAAAACAAAAATCGTGTAT 300
Db 241 TTAGCTTTTCACTAGCTTGTGGGGGTTTGACCCCTGAATAGCAAAACAAAAATCGTGTAT 300
QY 301 TTCAATGACGATTCATAGGTTTAAATTCGCGATCCCTGTAAACCCGAGCGACAGATTAGAA 360
Db 301 TTCAATGACGATTCATAGGTTTAAATTCGCGATTCCTGTAAACCCGAGCGACAGATTAGAA 360
QY 361 TACCAATTTAGAACTCTTAAAGCATAAGGGGCATGATCTGGCAAGTGGTGGCAGCGCTCAA 420
Db 361 TACCAATTTAGAACTCTTAAAGCATAAGGGGCATGATCTGGCAAGTGGTGGCAGCGCTCAA 420
QY 421 GTGGATGGCAAGTGGTGGCTGAAGCGGAATTTGAAGCCCATGATTGACAGAGAGATTAA 480
Db 421 GTGGATGGCAAGTGGTGGCTGAAGCGGAATTTGAAGCCCATGATCGCAGAGAGAGATTAA 480

RESULT 3
ABX66131
ID ABX66131 standard; DNA; 853 BP.
XX AC ABX66131;
XX DT 07-MAY-2003 (first entry)
XX DE Helicobacter pylori selected interacting domain (SID) DNA #730.
XX KW Protein-protein interaction; ulcer; selected interacting domain; SID;
XX KW gene; ds.
XX OS Helicobacter pylori.
XX PN WO200266501-A2.
XX PD 29-AUG-2002.
XX PF 28-DEC-2001; 2001WO-EP015428.
XX PR 02-JAN-2001; 2001US-0259302P.
XX PA (HYBR-) HYBRIGENICS.
XX PA (INSP) INST PASTEUR.
XX PI Legrain P, Rain J, Colland F, De Reuse H, Labigne A;
XX WPI; 2002-674910/72.
XX DR P-PSDB; ABU51387.
XX PT New complexes of protein-protein interactions in Helicobacter pylori,
XX PT useful for identifying modulating compounds for treating or preventing
XX PT ulcers in mammals.
XX PS Claim 7; Page 263; 642pp; English.
XX CC The invention describes a complex of protein-protein interactions in
XX CC Helicobacter pylori selected from 421 complexes given in the
XX CC specification. The complex of protein-protein interactions are useful for
XX CC screening for agents which modulate the interaction of proteins.
XX CC Modulating compounds which binds to a targeted bacterial protein may be
XX CC used for treating or preventing ulcers in a human or animal. This
XX CC sequence encodes a selected interacting domain (SID), identified via
XX CC protein-protein interactions
SQ Sequence 853 BP; 258 A; 163 C; 167 G; 265 T; 0 U; 0 Other;
Query Match 61.3%; Score 294; DB 6; Length 853;

Best Local Similarity 100.0%; Pred. No. 4.3e-80;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAACAAGCCCATCAAACTTGCAATCTCAATTTTTTATAGAGCATATCTTACAAATT 60
Db 560 ATGGAACAAGCCCATCAAACTTGCAATCTCAATTTTTTATAGAGCATATCTTACAAATT 619
QY 61 CTACTCTACCGCTATCCCATGCTTTTAGTGAGTAGAATATAGAGTTACAAGCCAATAAA 120
Db 620 CTACTCTACCGCTATCCCATGCTTTTAGTGAGTAGAATATAGAGTTACAAGCCAATAAA 679
QY 121 AAAATTGTCGCTTATAAGAATATCACATTTTAAATGAAGACGTGTTTAAACGGGCATTTCCCT 180
Db 680 AAAATTGTCGCTTATAAGAATATCACATTTTAAATGAAGACGTGTTTAAACGGGCATTTCCCT 739
QY 181 AATAAGCCCATTTTCCGGGCGTTTTCATCGTAGAGGGCATGCGCAACCGGAGGGTTT 240
Db 740 AATAAGCCCATTTTCCGGGCGTTTTCATCGTAGAGGGCATGCGCAACCGGAGGGTTT 799
QY 241 TTAGCCCTTCACTAGCTTGTGGGGGTTTGACCCCTGAATCGCAAAACAAAAATC 294
Db 800 TTAGCCCTTCACTAGCTTGTGGGGGTTTGACCCCTGAATCGCAAAACAAAAATC 853
RESULT 4
ACA30215
ID ACA30215 standard; DNA; 441 BP.
XX AC ACA30215;
XX DT 19-JUN-2003 (first entry)
XX DE Prokaryotic essential gene #11872.
XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;
XX KW drug design; gene.
XX OS Campylobacter jejuni.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX DR P-PSDB; ABU26345.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 14; SEQ ID NO 18085; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 441 BP; 152 A; 50 C; 101 G; 138 T; 0 U; 0 Other;

Query Match 28.0%; Score 134.4; DB 8; Length 441;
Best Local Similarity 58.4%; Pred. No. 6.6e-31;
Matches 256; Conservative 0; Mismatches 176; Indels 6; Gaps 1;
QY 39 TATAGAGCATATCTACAAATCTACTCCACCGCTATCCCATCTTTTAGTGATAGAT 98
DB 9 TCTATGCAATCTCAAGAAATCTTACCACCGTTATCTTTTACTAGTAGATAAAT 68
QY 99 TATAGAGTTACAAGCAATAAATAATGTCGTTATAGAATATACATTTTAAATGAAGA 158
DB 69 TACAGAAATTAAGATTAAGAGAGTGGTCTAGGATATAAATAACATCAGTATAGTGACCA 128
QY 159 CGTGTTTAACGGGCAATTCCTTAATAAGCCATTTTCCGGCGTTTTCATGCTAGAGGG 218
DB 129 TGTGTTTATGGGACATTTTCCAGGGCATCTATTTATCTCGAGTTTTCATTTAGAAG 188
QY 219 CATGGCGCAACCGGAGGGTGTTTTAGCCCTTCACTAGCTTGTGGGGTTTGACCCCTGAAT 278
DB 189 TATGCTCAACAGGTGGAGTTTACCTTTTGAAGTAT-----GGAAGATAAGTGGA 242
QY 279 CGCCAAAACAAAATCGTGATTTTCATGACGATTAAGGTTAAATTCGCATCCCTGT 338
DB 243 TCCAAAAGTAAAGTAGTTTATTTACAGGCATAGATGGAGCAAAATTTAGAAATCCTGT 302
QY 339 AACCCAGCGCAGATTAGATACCATTTTAGAGTCTTAAAGCATAGGCGCATGATCTG 398
DB 303 CGCTCTGGAGATAGGCTTGTATATGAATAGCGGTGTTAAATAATCGGTGAATATGTG 362
QY 399 GCAAGTGGGTGGCAGCGCTCAAGTGATGCGAAAGTGGTCTGAAGCGCAATTTGAAGC 458
DB 363 GATTTTAAAGGCGCAAGCTTTTGTAGATGGAATTTTAGTTGACAGGCGCGAGCTTAAGC 422
QY 459 CATGATTCAGAGAGAGA 476
DB 423 TATGATAGTGGATAATA 440

RESULT 5
ABX66187
ID ABX66187 standard; DNA; 659 BP.
XX
AC ABX66187;
XX
XX
DT 07-MAY-2003 (first entry)
XX
DE Helicobacter pylori selected interacting domain (SID) DNA #786.
XX

KW Protein-protein interaction; ulcer; selected interacting domain; SID;
KW gene; ds.
XX
OS Helicobacter pylori.
XX
PN WO200266501-A2.
XX
XX 29-AUG-2002.
XX
XX 28-DEC-2001; 2001WO-BP015428.
XX
XX 02-JAN-2003; 2001US-0259302P.
XX
XX (HYBR-) HYBRIGENICS.
XX (INSP) INST PASTEUR.
XX
PI Legrain P, Rain J, Colland F, De Reuse H, Labigne A;
XX WPI; 2002-674910/72.
XX P-PSDB; ABU51443.
XX
XX New complexes of protein-protein interactions in Helicobacter pylori,
XX useful for identifying modulating compounds for treating or preventing
XX ulcers in mammals.
XX
XX Claim 7; Page 277; 642pp; English.
XX
XX The invention describes a complex of protein-protein interactions in
XX Helicobacter pylori selected from 421 complexes given in the
XX specification. The complex of protein-protein interactions are useful for
XX screening for agents which modulate the interaction of proteins.
XX Modulating compounds which binds to a targeted bacterial protein may be
XX used for treating or preventing ulcers in a human or animal. This
XX sequence encodes a selected interacting domain (SID), identified via
XX protein-protein interactions
XX
SQ Sequence 659 BP; 200 A; 125 C; 123 G; 211 T; 0 U; 0 Other;
Query Match 20.8%; Score 100; DB 6; Length 659;
Best Local Similarity 100.0%; Pred. No. 3.3e-20;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAACAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTCAAAAT 60
DB 560 ATGGAACAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTCAAAAT 619
QY 61 CTACCTCACGCTATCCCATGCTTTTAGTGATAGAATTA 100
DB 620 CTACCTCACGCTATCCCATGCTTTTAGTGATAGAATTA 659
RESULT 6
ADS59846
ID ADS59846 standard; cDNA; 459 BP.
XX
XX ADS59846;
XX
XX 02-DEC-2004 (first entry)
XX
XX Bacterial polynucleotide #11833.
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polynucleotide; gene; ss.
XX
XX Bacteria.
XX
XX US2003233675-A1.
XX

PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.
 XX
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 3520; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 459 BP; 103 A; 141 C; 132 G; 83 T; 0 U; 0 Other;
 Query Match 20.7%; Score 99.2; DB 13; Length 459;
 Best Local Similarity 55.1%; Pred. No. 5e-20;
 Matches 225; Conservative 0; Mismatches 168; Indels 15; Gaps 1;
 QY 59 TTCTACCTACCGCTATCCATGCTTTTATAGTGATAGATATATAGAGTACAGCCAAATA 118
 DB 53 TCCTGCCACACCGCTATCCCTCTGATGATCGACGCGATCGACATCGATGCTGACG 112
 QY 119 AAAAAATTGCTCTATAGATATACATTTTAAATGAAGACGTTTAAACGGCATTTCC 178
 DB 113 ATTCCGCCATCGGCATCAAGAACGTCACATCAACGAGCGCGATTTTCAAGGGCATTTCC 172
 QY 179 CTAATAAGCCCATTTTCCGGCGGCTTTTCATCGTAGAGGCGATGGCGCAACGGGAGGGT 238
 DB 173 CGGAGCAGCGGTGATGCGGGCGTGTGATGTCGAGACCATGGCGCAGACGGGAGGG 232
 QY 239 TTTTACGCTTACTAGCTTGTGGGGTGTGACCTGAAATCGCCAAACAAAAAATCGTGT 298
 DB 233 CCATCTGCATCGCAGCTCGGGGGT-----CGAAGCGCTGCTGCTCT 277
 QY 299 ATTTCATGACGATGATAGAGTTAAATTCGGCATCGCTGTAAACCCAGCGGACAGATTAG 358
 DB 278 ATTTCTGACCATGACAAACGCGCAAAATTCGCGAAACCGGTGCTTCCGCGGCGACAGTTGA 337

QY 359 AATACATTTAGAGTCTTTAAGCATAGAGGCATGATCTGGCAAGTGGGTGGCACGGCTC 418
 DB 338 AGATTACGTCACAAAATAATCAAGAGCGCGCAACCTGCTCAAAATTCGCTGTGAAGCCC 397
 QY 419 AAGTGGATGGCAAAATGTCGCTGAAGCCGAATTTGAAAGCCATGATTG 466
 DB 398 TGGTCGATGGCAACCAAGGCGCGGAGATTTTCAGCCATGATGG 445
 RESULT 7
 ACA28687
 ID ACA28687 standard; DNA; 432 BP.
 XX
 AC ACA28687;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DB Prokaryotic essential gene #10344.
 XX
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX
 OS Clostridium botulinum.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR P-PSDB; ABU24817.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 16557; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 432 BP; 169 A; 43 C; 95 G; 125 T; 0 U; 0 Other;
 Query Match 19.8%; Score 95; DB 8; Length 432;
 Best Local Similarity 53.5%; Pred. No. 9.7e-19;
 Matches 234; Conservative 0; Mismatches 185; Indels 18; Gaps 1;
 QY 34 TTTTATAGAGCATCTTACAAATCTACCTCACCGTATCCCATGCTTTTAGTGGAT 93
 DB 13 TTAGATAAATGAATATAAAATATATCCGACAGATATCCATTTTATAGTAGAT 72
 QY 94 AGAATTATAGAGTTACAAGCCCAATAAAAAATTTGCGCTTATAAGAAATATCACTTTTAAT 153
 DB 73 AAAATACTGAATTAGAGAGGAGGAAAGCTGCGATGATATAAAAAATGTACAGCTAAT 132
 QY 154 GAAGAGCTTTTAAACGGGCAATTCCTATAAAGCCCATTTTCCGGCGGCTTTTGAATCGTA 213
 DB 133 GAATACCTTTTAAATGGGCAATTTCCGGAAGAACCAAGTAATGCTGGGCTTCTGATTATA 192
 QY 214 GAGGCGATGGCCAAACGGAGGGTTTGTAGCTTCACTAGCTTGTGGGGGTTTGACCT 273
 DB 193 GAAGCTTTAGCACAGGTGGAGCTGTGTATTTTAAAGTA-----AA 234
 QY 274 GAAATCGCCAAACAAAATCGTATTTTCATGACCATGTAAGGTATAAATCCGCATC 333
 DB 235 GAAGAGTTTAAAGGAAATAGCTTATTTGGAGGTATAAATAAGCTAAGTTCAGAAA 294
 QY 334 CCGTAAACCCAGGCGACAGATAGAAATACCAATTTAGAAAGTCTTAAAGCATAGGCGATG 393
 DB 295 AAGGTAGTACACAGGAGATGTTTAAACTTAGTATAGATCTTACTAAATAAAGGCGTT 354
 QY 394 ATCTGCAAGTGGTGGCAGGCTCAAGTGGATGCGCAAGTGTCTGCTGAAGCCGAATTG 453
 DB 355 GCAGGAGTAGGTAAAGCCCTAGCTACTGTAGATGGGAAAGTAGCTGCAAGAGCAGAATTA 414
 QY 454 AAAGCCATGATTCAGA 470
 DB 415 TTATTGTAAAGAAA 431

RESULT 8
 ACA33871
 ID ACA33871 standard; DNA; 417 BP.
 XX AC ACA33871;
 XX 19-JUN-2003 (first entry)
 DT Prokaryotic essential gene #15528.
 DE Prokaryotic essential gene;
 XX Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX Enterococcus faecium.
 OS WO200277183-A2.
 PN 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR P-PSDB; ABU30001.
 XX New antiseense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 21741; 1766pp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 417 BP; 147 A; 66 C; 99 G; 105 T; 0 U; 0 Other;
 Query Match 19.5%; Score 93.4; DB 8; Length 417;
 Best Local Similarity 53.6%; Pred. No. 3e-18;
 Matches 230; Conservative 0; Mismatches 181; Indels 18; Gaps 1;
 QY 40 ATAGAGCATATCTTACAAATTTACCTCACCGTATCCCATGCTTTTAGTGGATGAATT 99
 DB 7 ATACAGAAATATAAGAAATCATTTCCACATCGTTATCCATGCTTTTGCATGATCGATG 66
 QY 100 ATAGAGTTACAAGCCCAATAAAAAATTTGCTGCTTATAAGAAATATCACTTTTAATGAAGAC 159
 DB 67 GAAGAAATGGTTGAAGGAGAACCGATCGTTTCTAAGAAAAATGTAAACATCAATGAGGCC 126
 QY 160 GTGTTTAAACGGCAATTTCCCTATAAGCCCATTTTCCGGCGGCTTTTGCATCGTAGAGGC 219
 DB 127 TTTTTCAGAGACATTTTCCAGAGAACCTGTTATGCGAGGGGTATTGATTGTTGAAGCA 186
 QY 220 ATGGCGAAAAACGGAGGGGTTTTAGCCCTTCACTAGCTTTGTGGGGGTTTGACCCCTGAATC 279
 DB 187 ATGGCTCAAGCAGAGCAGTGTGCATTACTTTCT-----CTTGAACAA 228
 QY 280 GCCAAAAACAAAATCGTGTATTTCATGACGATTAAGAGTTAAATTCGCGATCCCTGTA 339
 DB 229 TTCAAGGGAAGACGGCTTATTTCGCGGATTTGGACAAAGCAAAATTCGTAAGAAAGTT 288

PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR P-PSDB; ABU32614.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 14; SEQ ID NO 24354; 1766pp; English.
 PS
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway;
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 435 BP; 142 A; 61 C; 104 G; 128 T; 0 U; 0 Other;

RESULT 11

ABQ69109
 ID ABQ69109 standard; DNA; 1755 BP.
 XX
 AC ABQ69109;
 XX
 DT 29-AUG-2003 (revised)
 DT 29-AUG-2002 (first entry)
 XX
 DE Listeria monocytogenes 4b contig DNA sequence #1875.
 XX
 KW Antibacterial; Listeria; food contamination; mutational analysis;
 KW infection; ds.
 XX
 OS Listeria monocytogenes ATCC 19115.
 XX
 PN WO200228891-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 04-OCT-2001; 2001WO-FR003061.
 XX
 PR 04-OCT-2000; 2000FR-00012697.
 XX
 XX (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 PI Kunst F, Glaser P;
 XX WPI; 2002-332479/37.
 DR
 XX New genomic sequences from Listeria species, useful for detection,
 PT treatment and prevention of infection, also related polypeptides,
 PT antibodies and modulators.
 XX
 PS Claim 14; SEQ ID NO 1922; 180pp; French.
 XX
 CC The present invention relates to nucleic acid sequences (ABQ67188-
 CC ABQ71212) from Listeria sp. The sequences are useful as probes and
 CC primers for identification and/or detection of Listeria (e.g. as
 CC contaminants in foods, or mutational analysis) and for analysis of gene
 CC expression. Proteins encoded by the nucleic acid sequences can be used to
 CC screen for compounds that modulate gene expression, replication and
 CC pathogenicity of Listeria (potential therapeutic agents), also for
 CC treating infections by Listeria, and are useful as immunogens in anti-
 CC Listeria vaccines. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. (Updated
 CC on 29-AUG-2003 to standardise OS field)
 XX
 SQ Sequence 1755 BP; 583 A; 238 C; 385 G; 548 T; 0 U; 1 Other;
 Query Match 18.5%; Score 88.6; DB 6; Length 1755;
 Best Local Similarity 64.3%; Pred. No. 1.7e-16;
 Matches 133; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
 QY 34 TTTTATAGACATATCTTACAAATTCCTACCGCTATCCCATGCTTTTAGTGAT 93
 DB 4 TTAGATATAAGAAATCAAGAGATTTTGGCTTCATCGTTATCCATTTTGTAGTAGAT 63
 QY 94 AGAATTATAGAGTTACAGCCCAATAAAAAATTTGCTTTATAAGAAATATCACTTTTAAT 153
 DB 64 AGAGTTATTTCTGTTGAAGAACCAAAAGATTTACCGCTATTAAAGATGTACAGCCAAT 123
 QY 154 GAAGACGTGTTTAAACGGCATTTCCCTAATAAGCCCATTTTCCCGGGCGTTTGTATCGTA 213
 DB 124 GAAGAAATTTTAAACGGGCACCTTCTCTAGTATCTCTGTAATGCCAGGAGTATTAAATAGTG 183
 QY 214 GAGGCGATGGCCCAACCGGAGGTTT 240
 DB 184 GAAGCTCTGGCCCAACCAAGTGAATT 210
 Query Match 18.5%; Score 88.6; DB 6; Length 1755;
 Best Local Similarity 64.3%; Pred. No. 1.7e-16;
 Matches 133; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
 QY 34 TTTTATAGACATATCTTACAAATTCCTACCGCTATCCCATGCTTTTAGTGAT 93
 DB 830 TTAGATATAAGAAATCAAGAGATTTTGGCTTCATCGTTATCCATTTTGTAGTAGAT 889
 QY 94 AGAATTATAGAGTTACAGCCCAATAAAAAATTTGCTTTATAAGAAATATCACTTTTAAT 153
 DB 890 AGAGTTATTTCTGTTGAAGAACCAAAAGATTTACCGCTATTAAAGATGTACAGCCAAT 949
 QY 154 GAAGACGTGTTTAAACGGCATTTCCCTAATAAGCCCATTTTCCCGGGCGTTTGTATCGTA 213
 DB 950 GAAGAAATTTTAAACGGGCACCTTCTCTAGTATCTCTGTAATGCCAGGAGTATTAAATAGTG 1009
 QY 214 GAGGCGATGGCCCAACCGGAGGTTT 240
 DB 1010 GAAGCTCTAGCCCAACCAAGTGAATT 1036
 RESULT 12

ABQ70710
 ID ABQ70710 standard; DNA; 1927 BP.
 XX AC ABQ70710;
 XX XX
 DT 29-AUG-2003 (revised)
 XX DT 29-AUG-2002 (first entry)
 XX XX
 DE Listeria monocytogenes 4b contig DNA sequence #652.
 XX XX
 KW Antibacterial; Listeria; food contamination; mutational analysis;
 XX infection; ds.
 XX XX
 OS Listeria monocytogenes ATCC 19115.
 XX XX
 PN WO200228891-A2.
 XX XX
 PD 11-APR-2002.
 XX XX
 PF 04-OCT-2001; 2001WO-FR003061.
 XX XX
 PR 04-OCT-2000; 2000FR-00012697.
 XX XX
 PA (INSP) INST PASTEUR.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX PI Kunst F, Glaser P;
 XX WPI; 2002-332479/37.
 DR XX
 PT New genomic sequences from Listeria species, useful for detection,
 PT treatment and prevention of infection, also related polypeptides,
 PT antibodies and modulators.
 XX XX
 PS Claim 14; SEQ ID NO 3523; 180pp; French.
 XX XX
 CC The present invention relates to nucleic acid sequences (ABQ67188-
 CC ABQ71212) from Listeria sp. The sequences are useful as probes and
 CC primers for identification and/or detection of Listeria (e.g. as
 CC contaminants in foods, or mutational analysis) and for analysis of gene
 CC expression. Proteins encoded by the nucleic acid sequences can be used to
 CC screen for compounds that modulate gene expression, replication and
 CC pathogenicity of Listeria (potential therapeutic agents), also for
 CC treating infections by Listeria, and are useful as immunogens in anti-
 CC Listeria vaccines. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. (Updated
 CC on 29-AUG-2003 to standardise OS field)
 XX XX
 SQ Sequence 1927 BP; 641 A; 264 C; 426 G; 596 T; 0 U; 0 Other;
 Query Match 18.5%; Score 88.6; DB 6; Length 1927;
 Best Local Similarity 64.3%; Pred. No. 1.8e-16;
 Matches 133; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
 QY 34 TTTTATATAGCATATCTTACAAATTTCTACCTCACCGCTATCCCATGCTTTTAGTGAT 93
 Db 990 TTAGATATTAAAGAAATCAAGAGATTTTGCTCATCGTTATCCATTTTGTAGTAGT 1049
 QY 94 AGAATTATAGAGTTTACAGCCCAATAAAAAATTTGCTTTATAAGAAATATCACTTTTAAT 153
 Db 1050 AGAGTTATTCTGTTGAAGAGGCAAAAAGTTTACCGCTATTAAAGAAATGTAACAGCAAT 1109
 QY 154 GAAGAGCTGTTTAAACGGGCAATTCCTTAATAGCCCAATTTCCCGGGCGTTTGTATCGTA 213
 Db 1110 GAAGAATTTTTTAAACGGGCACCTTCTCGTAGTATCCTGTAATGCCAGGAGTTTAAATAGTG 1169
 QY 214 GAGGCGATCGCGCAACAGCGAGGTTT 240
 Db 1170 GAAGCTCTAGCGCAAACTAGTGAATTT 1196
 RESULT 13

ABA03041_25/c
 Continuation (26 of 30) of ABA03041 from base 2500001 (Listeria monocytogenes EGD-e genom
 WP Sequence split into 30 fragments LOCUS ABA03041 Accession ABA03041
 WP Fragment Name Begin End
 WP ABA03041_00 1 110000
 WP ABA03041_01 100001 210000
 WP ABA03041_02 200001 310000
 WP ABA03041_03 300001 410000
 WP ABA03041_04 400001 510000
 WP ABA03041_05 500001 610000
 WP ABA03041_06 600001 710000
 WP ABA03041_07 700001 810000
 WP ABA03041_08 800001 910000
 WP ABA03041_09 900001 1010000
 WP ABA03041_10 1000001 1110000
 WP ABA03041_11 1100001 1210000
 WP ABA03041_12 1200001 1310000
 WP ABA03041_13 1300001 1410000
 WP ABA03041_14 1400001 1510000
 WP ABA03041_15 1500001 1610000
 WP ABA03041_16 1600001 1710000
 WP ABA03041_17 1700001 1810000
 WP ABA03041_18 1800001 1910000
 WP ABA03041_19 1900001 2010000
 WP ABA03041_20 2000001 2110000
 WP ABA03041_21 2100001 2210000
 WP ABA03041_22 2200001 2310000
 WP ABA03041_23 2300001 2410000
 WP ABA03041_24 2400001 2510000
 WP ABA03041_25 2500001 2610000
 WP ABA03041_26 2600001 2710000
 WP ABA03041_27 2700001 2810000
 WP ABA03041_28 2800001 2910000
 WP ABA03041_29 2900001 2944528
 Query Match 18.5%; Score 88.6; DB 6; Length 110000;
 Best Local Similarity 64.3%; Pred. No. 1.1e-15;
 Matches 133; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
 QY 34 TTTTATATAGCATATCTTACAAATTTCTACCTCACCGCTATCCCATGCTTTTAGTGAT 93
 Db 102680 TTAGATATTAAAGAAATCAAGAGATTTTGCTCATCGTTATCCATTTTGTAGTAGT 102621
 QY 94 AGAATTATAGAGTTTACAGCCCAATAAAAAATTTGCTTTATAAGAAATATCACTTTTAAT 153
 Db 102620 AGAGTTATTCTGTTGAAGAGGCAAAAAGTTTACCGCTATTAAAGAAATGTAACAGCCAT 102561
 QY 154 GAAGAGCTGTTTAAACGGGCAATTTCCCTTAATAAGCCCAATTTTCCCGGGCGTTTGTATCGTA 213
 Db 102560 GAAGAATTTTTTAAACGGGCACCTTCTCTGAGTATCCTGTAATGCCAGGAGTATTAAATAGTG 102501
 QY 214 GAGGCGATCGCGCAACAGCGAGGTTT 240
 Db 102500 GAAGCTCTGCGCAACAACTAGTGAATTT 102474
 RESULT 14
 ABA03041_26/c
 Continuation (27 of 30) of ABA03041 from base 2600001 (Listeria monocytogenes EGD-e genom
 WP Sequence split into 30 fragments LOCUS ABA03041 Accession ABA03041
 WP Fragment Name Begin End
 WP ABA03041_00 1 110000
 WP ABA03041_01 100001 210000
 WP ABA03041_02 200001 310000
 WP ABA03041_03 300001 410000
 WP ABA03041_04 400001 510000
 WP ABA03041_05 500001 610000
 WP ABA03041_06 600001 710000
 WP ABA03041_07 700001 810000
 WP ABA03041_08 800001 910000
 WP ABA03041_09 900001 1010000
 WP ABA03041_10 1000001 1110000
 WP ABA03041_11 1100001 1210000

WP	ABA03041_12	1200001	1310000
WP	ABA03041_13	1300001	1410000
WP	ABA03041_14	1400001	1510000
WP	ABA03041_15	1500001	1610000
WP	ABA03041_16	1600001	1710000
WP	ABA03041_17	1700001	1810000
WP	ABA03041_18	1800001	1910000
WP	ABA03041_19	1900001	2010000
WP	ABA03041_20	2000001	2110000
WP	ABA03041_21	2100001	2210000
WP	ABA03041_22	2200001	2310000
WP	ABA03041_23	2300001	2410000
WP	ABA03041_24	2400001	2510000
WP	ABA03041_25	2500001	2610000
WP	ABA03041_26	2600001	2710000
WP	ABA03041_27	2700001	2810000
WP	ABA03041_28	2800001	2910000
WP	ABA03041_29	2900001	2944528

Query Match 18.5%; Score 88.6; DB 6; Length 110000;
Best Local Similarity 64.3%; Pred. No. 1.e-15;
Matches 133; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy	34	TTTTTTATAGACGATCTTACAATAATTCTACCTCACGCTATCCATGCTTTTAGTGGAAT	93
Db	2680	TTAGATATTAAAGAAAATCAAAGAGATTTCGCCTCATCGTTATCCATTTTGTGGTAGTAGAT	2621
Qy	94	AGRATTATAGAGTTTACAAAGCCAATAAAAAAATTCGCTTATAAGAATATCACCITTTTAAT	153
Db	2620	AGAGTTATTCTGTTGAAGAAGGCCAAAAGTTAACCGCTATTAAAGAATGTAAACAGCCAAAT	2561
Qy	154	GAAGACGCTGTTTAAACGGGCGATTTCCCTTAATAAGCCCAATTTTCCCGGGCGTTTGTATCGTA	213
Db	2560	GAAGAATTTTAAACGGGCGACTTTCTCTGAGTATCCTGTAAATGCCAGGAGTATTAAATAGTG	2501
Qy	214	GAGGGCATGGCGCAACGAGGAGGTTT	240
Db	2500	GAAGCTCTGGCGCAACAAGTGGAAAT	2474

RESULT 15
ACA43412
ID ACA43412 standard; DNA; 459 BP.
XX AC ACA43412;
XX AC
DT 19-JUN-2003 (first entry)
XX DE
XX DE Prokaryotic essential gene #25069.
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX OS Pasteurella multocida.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
DR WPI: 2003-029926/02.

P-PSDB; ABU39542.

XX New antiseNSE nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX
XX
PS Claim 14; SEQ ID NO 31282; 1766pp; English.

XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antiseNSE sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid;
CC encoding a polypeptide whose expression is inhibited by the antiseNSE
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antiseNSE nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antiseNSE nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 459 BP; 130 A; 82 C; 110 G; 137 T; 0 U; 0 Other;

Query Match 18.4%; Score 88.2; DB 8; Length 459;
Best Local Similarity 53.3%; Pred. NO. 1.3e-16;
Matches 220; Conservative 0; Mismatches 178; Indels 15; Gaps 1

QY 42 AGAGCATATCTTACAAATTCCTACCGTATCCCATGCTTTTGTAGTCGATAGAAATTAT 101
DB 39 AAATGAAATTATGAAATTCCTTCCTCACCGTATCCTTTTATTAGTGGATCGTGTGT 98

QY 102 AGAGTTACAAGCCCAATAAAAAATTCGCTGTATAGAAATATCACATTTTAATGAAGACGT 161
DB 99 GGAATTATCAAGAGTAAATGCTTTAAGCTGTAAAAATATTAGCGTAATGAGCCTTG 151

QY 162 GTTTTAAAGGGCAATTTCCCTAATAAGCCCATTTTCCGGCGGTTTTCATCGTAGAGGCAT 221
DB 159 CTTTTCAGGGAATTTCCAGAACACACCGATTTTTCGGGTGTGTTCATTTTGGAAAGCAT 211

QY 222 GGCGCAAAACGGAGGGTTTTTAGCCTTCACCTAGCTTGTGGGGGTTTGAACCTGAATCGC 281
DB 219 GGCAACAGCGACGGGGTGCTCGCCTGTAAACTTTACCGTCAATTAG----- 261

QY 282 CAAACAAAAATCGTGATTTTCATGACGATTCATAGAGGTTTAAATTCGCGATCCCTGTAA 341
DB 266 --AAATGAAATTTATTTTGTGCAATTCGATTAATGCGCGCTTTAAACGTCCTGTGTT 321

QY 342 CCCAGGCGACAGATTAGAAATACCAATTTAGAAGCTTTTAAAGCAATAAGGCAATGATCTGGCA 401
DB 324 GCCTGGTGCATCAATGGTGTAGAAAGTTTCACCTCTTCTTGAAGAGCGTCGCGGTATTACGCG 381

QY 402 AGTGGGTGGCAGCGCTCAAGTGGATGCGAAAGTGGTCGCTGAAGCCGAATTGA 454
DB 384 TTTTCACTGGTGGCAACCGCTAGATGGTCAAGTGGTGTGCGAAGCAGAAATTAA 436

Search completed: January 12, 2006, 01:29:30
Job time : 464 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 01:07:09 ; Search time 3071 Seconds

(without alignments)
7312.862 Million cell updates/sec

Title: US-10-662-126-36

Perfect score: 480

Sequence: 1 atggcaaaagcatcaaaaa.....tgattcgagagagattaa 480

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_hic:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_est7:*

9: gb_gsa1:*

10: gb_gsa2:*

11: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	85	17.7	724	6	CD830511	CD830511 BN40.045M
2	84.8	17.7	851	8	DR926513	DR926513 EST111805
3	84.4	17.6	565	5	BQ577106	BQ577106 PESToabl
4	84	17.5	845	4	CNS0092WL	EX830012 Arabidops
5	83.6	17.4	939	9	BZ567344	BZ567344 pacs2-164
6	83.2	17.3	820	8	DR948072	DR948072 EST113961
7	83.2	17.3	830	8	DR942398	DR942398 EST113393
8	83.2	17.3	841	8	DR938067	DR938067 EST112960
9	83.2	17.3	848	8	DR929905	DR929905 EST112144
10	83.2	17.3	877	8	DR931821	DR931821 EST112336
11	83	17.3	465	1	AJ750880	AJ750880 AJ750880
12	82.8	17.2	631	8	DN910931	DN910931 Cuc6_3 Cl
13	82.4	17.2	583	1	AI994492	AI994492 70149847
14	82.4	17.2	598	1	AU235591	AU235591 AU235591
15	82.4	17.2	774	4	CNS0A1U5	EX834110 Arabidops
16	82.4	17.2	799	4	CNS00929N	EX833007 Arabidops
17	82.4	17.2	840	4	CNS0A1T2	EX833404 Arabidops
18	82	17.1	848	8	DR938068	DR938068 EST112960
19	81.8	17.0	764	7	CN734884	CN734884 25RDNM_U
20	81.6	17.0	693	8	CK045132	CK045132 UCRCS07_2
21	80.8	16.8	457	5	BQ791261	BQ791261 E4137 CH1
22	80.8	16.8	471	2	BG544298	BG544298 E2138 CH1

23	80.6	16.8	538	7	CK937166	CK937166 CGF100451
24	80.6	16.8	728	8	CX542609	CX542609 UCRPT01_5
c 25	80.6	16.8	791	8	CV887178	CV887178 UCRCS04_2
c 26	80.6	16.8	845	7	CF832673	CF832673 UCRCS02_0
27	80.2	16.7	616	6	CD822254	CD822254 BN25.0441
28	80.2	16.7	819	8	DR948073	DR948073 EST113961
29	80.2	16.7	852	8	DR929906	DR929906 EST112144
30	79.8	16.6	854	8	DR925940	DR925940 EST111747
c 31	79.6	16.6	629	6	CF515461	CF515461 CAP0001_1
c 32	79.6	16.6	703	6	CB971995	CB971995 CAB10006_
c 33	79.6	16.6	747	6	CB972347	CB972347 CAB10006_
c 34	79.6	16.6	760	6	CB972067	CB972067 CAB10006_
c 35	79.6	16.6	766	6	CB972267	CB972267 CAB10006_
c 36	79.6	16.6	772	6	CF210321	CF210321 CAB20006_
c 37	79.6	16.6	786	5	BQ796669	BQ796669 EST 5607
38	79.6	16.6	870	6	CF515372	CF515372 CAP0001_1
39	79.6	16.6	885	6	CF210235	CF210235 CAB20006_
40	79.2	16.5	547	8	DN776664	DN776664 E4179 [C6
41	79.2	16.5	586	3	BP530330	BP530330 BP530330
42	79	16.5	515	7	CO908062	CO908062 BU0202650
43	79	16.5	539	8	CV882970	CV882970 Mdfit3116
44	79	16.5	570	8	CV881159	CV881159 Mdfit3113
45	79	16.5	683	7	CN181917	CN181917 UCRCS04_0

ALIGNMENTS

RESULT 1
LOCUS CD830511 724 bp mRNA linear EST 10-JUL-2003
DEFINITION BN40.045M05F011228 BN40 Brassica napus cDNA clone BN40045M05, mRNA
sequence.
ACCESSION CD830511
VERSION CD830511.1 GI:32512451
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 724)
AUTHORS Genoplatte.
TITLE Genoplatte, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplatte
Genoplatte
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplatte' (<http://www.genoplatte.com>)
and <http://genoplatte-info.inbioigen.fr>.

FEATURES
source Location/Qualifiers
1..724
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jet Neuf"
/db_xref="taxon:3708"
/clone="BN40045M05"
/tissue_type="seed"
/clone_lib="BN40"

ORIGIN
Query Match 17.7%; Score 85; DB 6; Length 724;
Best Local Similarity 64.5%; Pred. No. 1.3e-13;
Matches 127; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 40 ATAGACATCTTACAAATCTTACCTACCGTATCCCATGCTTTAGTGATAGTAATT 99
DB 303 ATTAACAGATACGAGAGATTTTGCCTTCACAGGTTCCCGTTCTGTAGTGATAGTAGT 362
QY 100 ATAGAGTTACAAGCCAATAAAAAAATTTGCGCTTATAAGAAATATCACTTTTAAATGAAGAC 159

Db 363 ATAGAGTACACAGCTGGTGAATCTGGGTAGCTATCAAGAACGTTTACCATTATGCAAT 422
 QY 160 GTGTTTAAAGCGGCAATTCCTTAATAGCCCAATTTCCCGGGCGTTTGTGATCGTAGAGGGC 219
 Db 423 TTCTTTCCCTGGGCAATTTCCCGAGAGGCCCATATGCTTGGTGTCTCATGTTGAGGCC 482
 QY 220 ATGGCGCAACACGGAGG 236
 Db 483 ATGGCTCAGGTGGAGG 499

RESULT 2
 LOCUS DR926513/c
 DEFINITION EST1118052 Aquilegia cDNA library Aquilegia formosa x Aquilegia pubescens cDNA clone COINE67, mRNA sequence.
 ACCESSION DR926513
 VERSION DR926513.1 GI:71695876
 SOURCE EST.
 ORGANISM Aquilegia formosa x Aquilegia pubescens
 Aquilegia formosa x Aquilegia pubescens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
 Ranunculaceae; Aquilegia.
 REFERENCE 1 (bases 1 to 851)
 AUTHORS Hodges, S.A., Rensink, W., Buell, C.R., Borevitz, J., Kramer, E., Nordborg, M., and Tomkins, J.
 TITLE Generation of ESTs from Aquilegia
 JOURNAL Unpublished (2005)
 COMMENT Other ESTs: EST1118053
 Contact: Scott Hodges
 Department of Ecology, Evolution and Marine Biology
 University of California, Santa Barbara
 Santa Barbara, CA 93106, USA
 Tel: 805 893 7813
 Fax: 805 893 4724
 Email: hodges@lifesci.ucsb.edu
 Seq primer: TTTTCTTTTCTTTTCTTTTCTTTT (where N = A, G & C).

FEATURES
 source
 1..851
 /organism="Aquilegia formosa x Aquilegia pubescens"
 /mol_type="mRNA"
 /db_xref="taxon:338618"
 /clone="COINE67"
 /tissue_type="mixed shoot and floral apical meristems, flower buds, leaves and roots"
 /lab_host="DH10B T1 (T1 and T5 phage resistance)"
 /clone_lib="Aquilegia cDNA library"
 /notes="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI; F2, F3, and F4 lines of Aquilegia formosa X A. pubescens were grown from seed in greenhouses at UC Santa Barbara. From these plants three sets of tissue were collected: 1) Small flower buds (<10 mm) and very young inflorescences (71 & 29% by weight respectively), 2) Medium (7-20 mm) and large (at or near anthesis) flower buds (65 & 35% by weight respectively) and 3) Shoot apical meristems. A fourth set of tissue was collected from plants of A. formosa. These plants were grown from seed in sand and at approximately 1 month root tissue and leaf tissue of various developmental stages were collected (84 & 16% by weight respectively). Total RNA was extracted from each set of tissue and pooled in the following proportions: 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled total RNA, mRNA was extracted and enriched for full-length messages and then normalized with proprietary methods by Invitrogen."

ORIGIN

Query Match 17.7%; Score 84.8; DB 8; Length 851;
 Best Local Similarity 50.6%; Pred. No. 1.5e-13;
 Matches 243; Conservative 0; Mismatches 222; Indels 15; Gaps 1;

QY 1 ATGGAAACAAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
 Db 640 ATTGAAAAGAAAGATTCAAGAGCGTTTCCAAACAATAATGGATATTGAAAAGATTTCGTAACATT 581
 QY 61 CTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTTACAAGCCAATAAA 120
 Db 580 TTGCCTCATAGGTTTCGTTTCTTAGTTGATAGAGTGAATGATATGATCCTCGGAGTT 521
 QY 121 AAAAATTGTGCTTTATAGAATATCACTTTTAAATGAAGACGCTGTTTAAACGGGCATTTCCCT 180
 Db 520 TCGGCGGTTGCTATTAAAGATGTTACGATTAAATGATAATTTCTTTCCAGGCCATTTTCTCT 461
 QY 181 AATAAGCCCATTTTCCCGGCGGTTTTCATCGTAGAGGCATGGCGCAACGGGAGGGTTT 240
 Db 460 GAGAGGCCAATTATGCTGCTGTTCTTATGATTGAGGCAATGGCTCAAGTTGGAGGCTTG 401
 QY 241 TTAGCCTTCACTAGCTTGTGCGGGGTTTGACCTGAAATCGCCAAACAAAAAATCGTGAT 300
 Db 400 GTAATGCTGACCCAGAGTGGGTGTTCTCGTG-----AAAAATTTCTTC 356
 QY 301 TTCAATGACCAATTGATAAGGTTAAATTCGCGATCCCTGTAACCCAGGCGACAGATTAGAA 360
 Db 355 TTACAGGAATTGACAAAGTCAGATTCGGAACCTGTGGTTGCAGGGGATACGCTTGTA 296
 QY 361 TACCATTTAGAGTCTTTAAAGCATATAGGGCATGATCTGGCAAGTGGGTGGCAGCGCTCAA 420
 Db 295 ATGAGAATGATTCTCATATAAGAAACGCTTTGGAGTAGTTAAGATGGAGGGGAAGCGCTAT 236
 QY 421 GTGGATGGCAAGTGGTCTGCTGAAGCCGAATTGAAGCCATGATTGCAGAGAGAGATTAA 480
 Db 235 TTTGGGGGTGAGGTGGTATGTGACGGAGAGTTACGATGGCTCTTTGGGAAGAGCGAATGA 176

RESULT 3
 LOCUS BQ577106
 DEFINITION PFESToab13h01.y1 Plasmodium falciparum 3D7 asexual cDNA Plasmodium falciparum 3D7 cDNA 5' similar to TR:O97376 O97376
 BETA-HYDROXYACYL-ACP DEHYDRATASE PRECURSOR. i, mRNA sequence.

ACCESSION BQ577106
 VERSION BQ577106.1 GI:21480423
 KEYWORDS EST.
 SOURCE Plasmodium falciparum 3D7
 ORGANISM Plasmodium falciparum 3D7
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 REFERENCE 1 (bases 1 to 565)
 AUTHORS Tang, K., Cole, R., Chakrabarti, D., Haywood, R., Clifton, S., Pape, D., Marr, M., Hillier, L., Martin, J., Wylie, F., Dante, M., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Jentes, E., Ronko, I., Teagareishvili, R., Belaygorod, L., Franklin, C., Carr, L., Grow, A., Maguire, L., Richey, J., Wadkins, J., Kennedy, S., Levinso, D., Waterston, R., Wilson, R. and Sibley, D.

TITLE WashU Plasmodium EST Project
 JOURNAL Unpublished (2001)
 COMMENT Contact: L. David Sibley
 WashU Plasmodium EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library was constructed by Debopam Chakrabarti DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: L. David Sibley (sibley@borcim.wustl.edu), Washington University
 Seq primer: -40UP from Gibco
 High quality sequence stop: 423.

FEATURES

source
 1..565
 /organism="Plasmodium falciparum 3D7"
 /mol_type="mRNA"
 /db_xref="taxon:36329"
 /lab_host="DH10B (GeneHog, Invitrogen, Inc.)"

source
1...939
/organism="Pseudomonas aeruginosa"
/mol_type="Genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pac2-164_6892"
/clone_lib="pac2-164"
/note="clinical isolate 2-164 Whole genomic shotgun library."

ORIGIN

Query Match 17.4%; Score 83.6; DB 9; Length 939;
Best Local Similarity 53.9%; Pred. No. 3.5e-13;
Matches 223; Conservative 0; Mismatches 179; Indels 12; Gaps 2;
QY 44 AGCATATCTTACAAATTCACCGTATCCCATGCTTTTAGTGGATAGATATATAG 103
DB 65 AGCTTGATATCAATTCCTGACGCCCGCTATCCATCTCTGCTGGATCGGTGTAG 124
QY 104 AGTTCAAGCCAA---TAAAAAATTTGCTGTATTAAGAAATATCACTTTTAATGAAGACG 160
DB 125 AGCTGGACATCAGGAGGCAAGCGCATTCGCGCTTACAAGAAATGTCAATGAGCCGT 184
QY 161 TGTTTAAACGGGATTTCCCTAATAGCCCATTTTCCGGCGGTTTGTATCGTAGAGGCA 220
DB 185 TCTTCAACGGACATTTCCCGGAGCACCCGATCATGCGGGCGGTGTGATCATCGAGGCGA 244
QY 221 TGGCGCAACAGGGGGTTTTATAGCTTCCATAGCTTGTGGGGTTTGACCTTGAATCG 280
DB 245 TGGCCAGCGCGCGGTATCTCGGTTTCA-----AGATGCTCGACGTGAAGCCCG 295
QY 281 CCAAAACAAAATTCGTGTATTTTCATGACGATTTGATAGGTTAAATTCGCGATCCCTGTAA 340
DB 296 CGACGGCACCTTTACTACTTGTGCTCGGACAGCTGCGCTTCCGCGACCGGTCC 355
QY 341 CCCCAGGCGACAGATTAGAAATCAATTTAGAGTCTTAAAGCATAAGGCGATGATCTGGC 400
DB 356 TGGCGGGGACCAACTGCACTGACGCGCAAGTTTATCAGCGTGAAGCGCAGCATCTGGA 415
QY 401 AGTGGGTGGCAGCGCTCAAGTGGATGGCAAGTGTCTGCTGAAGCGGAATGA 454
DB 416 AATTGACTGCCACGCTACCGTGCAGCAGCAAGCCGTTATGCTCGGCTGAATCA 469

RESULT 6
DR948072/c
LOCUS
DEFINITION
EST1139611 Aquilegia cDNA library Aquilegia formosa x Aquilegia pubescens cDNA clone CO1R896, mRNA sequence.
ACCESSION
DR948072
VERSION
DR948072.1 GI:71717435
KEYWORDS
EST.
SOURCE
Aquilegia formosa x Aquilegia pubescens
ORGANISM
Aquilegia formosa x Aquilegia pubescens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales; Ranunculaceae; Aquilegia.
REFERENCE
1 (bases 1 to 820)
Hodges, S.A., Rensink, W., Buell, C.R., Borevitz, J., Kramer, E., Nordborg, M. and Tomkins, J.
Generation of ESTs from Aquilegia
Unpublished (2005)
Other ESTs: EST1139612
Contact: Scott Hodges
Department of Ecology, Evolution and Marine Biology
University of California, Santa Barbara
Santa Barbara, CA 93106, USA
Tel: 805 893 7813
Fax: 805 893 4724
Email: hodges@lifesci.ucsb.edu
Seq primer: TTTTCTTTTCTTTTCTTTTCTTTT (where N = A, G & C).
Location/Qualifiers
1..820

/organism="Aquilegia formosa x Aquilegia pubescens"
/mol_type="mRNA"
/db_xref="taxon:338618"
/clone="CO1R896"
/tissue type="mixed shoot and floral apical meristems, flower buds, leaves and roots"
/lab_host="DH10B T1 (T1 and T5 phage resistance)"
/clone_lib="Aquilegia cDNA library"
/note="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI; F2, F3, and F4 lines of Aquilegia formosa x A. pubescens were grown from seed in greenhouses at UC Santa Barbara. From these plants three sets of tissue were collected: 1) Small flower buds (<10 mm) and very young inflorescences (71 & 29% by weight respectively), 2) Medium (7-20 mm) and large (at or near anthesis) flower buds (65 & 35% by weight respectively) and 3) Shoot apical meristems. A fourth set of tissue was collected from plants of A. formosa. These plants were grown from seed in sand and at approximately 1 month root tissue and leaf tissue of various developmental stages were collected (84 & 16% by weight respectively). Total RNA was extracted from each set of tissue and pooled in the following proportions: 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled total RNA, mRNA was extracted and enriched for full-length messages and then normalized with proprietary methods by Invitrogen."

ORIGIN

Query Match 17.3%; Score 83.2; DB 8; Length 820;
Best Local Similarity 50.4%; Pred. No. 4.4e-13;
Matches 242; Conservative 0; Mismatches 223; Indels 15; Gaps 1;
QY 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGACATATCTTACAATT 60
DB 641 ATTGAAAAGAGATTCAAGCGTTTCCACAATAATGATATGAAAGATTGTAACATT 582
QY 61 CTACCTCACCGCTATCCATGCTTTTAGTGGATAGAAATTATAGAGTTACAGGCAATAA 120
DB 581 TTGCCTCATAGGTTTCCGTTTCTTAGTTGATAGAGTGATGAATATGATCTCGAGTT 522
QY 121 AATATGTCGCTTATAGATATACCTTTTAAAGACAGCTTTTAAACGGCATTTCCCT 180
DB 521 TCGCGGTTGCTATTAAAGAAATGTACGATTAATGATAATTTCTTCCAGGCCATTTCT 462
QY 181 AATAAGCCCATTTTCCCGGCGTTTTCATGCTAGAGGCGATGCGCAACCGGAGGTTT 240
DB 461 GAGAGCCCAATTATGCTGCTGTTCTTATGATGAGCAATGCTCAAGTTGGAGCTTG 402
QY 241 TTAGCCTTCACTAGCTTTGCGGGGTTTGACCCCTGAAATCGCAAAACAAAATCGTGTAT 300
DB 401 GTAAATGCTGACCCCGAAGTGGGTGTTCTCGTG-----AAAAATTTCTTC 357
QY 301 TTATGACAGATTGATAGGTTAAATTCGCTATCCGATCCCTTAACCCCGGACAGATTAGAA 360
DB 356 TTATACAGGAATTGACAAAGTCAGATTCGGAACCTGTTGTCAGGGGATACGCTTGTGA 297
QY 361 TACCATTTTAAAGTCTTTAAAGCATTAAGGCGATGATCTGGCAAGTGGGTGGCCAGCTCAA 420
DB 296 ATGAGAAATGATTTCTATTAAGAAACCGCTTTGGAGTAGTTAGATGGAGGGAAGGCCTAT 237
QY 421 GTGGATGGCAAGTGGTTCGCTGAAGCCGAAATTGAAGCCATGATTGACAGAGAGATTAA 480
DB 236 GTTGGGGTGAAGTGGTATGTGACGGAGAGTTTACGATGGCTCTTTGGGAAGCGAATGA 177

RESULT 7
DR942398/c
LOCUS
DEFINITION
EST1133937 Aquilegia cDNA library Aquilegia formosa x Aquilegia pubescens cDNA clone CO1QA71, mRNA sequence.
ACCESSION
DR942398
VERSION
DR942398.1 GI:71711761
KEYWORDS
EST.

SOURCE Aquilegia formosa x Aquilegia pubescens
 ORGANISM Aquilegia formosa x Aquilegia pubescens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
 Ranunculaceae; Aquilegia.
 REFERENCE 1 (bases 1 to 830)
 AUTHORS Hodges S.A., Rensink W., Buell C.R., Borevitz J., Kramer E.,
 Nordborg M. and Tomkins J.
 TITLE Generation of ESTs from Aquilegia
 JOURNAL Unpublished (2005)
 COMMENT Other ESTs: EST1133938
 Contact: Scott Hodges
 Department of Ecology, Evolution and Marine Biology
 University of California, Santa Barbara
 Santa Barbara, CA 93106, USA
 Tel: 805 893 7813
 Fax: 805 893 4724
 Email: hodges@lifesci.ucsb.edu
 Seq primer: TTTTATTTTATTTTATTTTATTTT (where N = A, G & C).
 FEATURES
 source
 1..830
 /organism="Aquilegia formosa x Aquilegia pubescens"
 /mol_type="mRNA"
 /db_xref="taxon:338618"
 /clone="CO10A71"
 /tissue_type="mixed shoot and floral apical meristems,
 flower buds, leaves and roots"
 /lab_host="DH10B T1 (T1 and T5 phage resistance)"
 /clone_lib="Aquilegia cDNA library"
 /note="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI;
 F2, F3, and F4 lines of Aquilegia formosa x A. pubescens
 were grown from seed in greenhouses at UC Santa Barbara.
 From these plants three sets of tissue were collected: 1)
 Small flower buds (<10 mm) and very young inflorescences
 (71 & 29% by weight respectively), 2) Medium (7-20 mm) and
 large (at or near anthesis) flower buds (65 & 35% by
 weight respectively) and 3) Shoot apical meristems. A
 fourth set of tissue was collected from plants of A.
 formosa. These plants were grown from seed in sand and at
 approximately 1 month root tissue and leaf tissue of
 various developmental stages were collected (84 & 16% by
 weight respectively). Total RNA was extracted from each
 set of tissue and pooled in the following proportions:
 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled
 total RNA, mRNA was extracted and enriched for full-length
 messages and then normalized with proprietary methods by
 Invitrogen."
 ORIGIN
 Query Match 17.3%; Score 83.2; DB 8; Length 830;
 Best Local Similarity 50.4%; Pred. No. 4.5e-13;
 Matches 242; Conservative 0; Mismatches 223; Indels 15; Gaps 1;
 QY 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGATATCTTCAAAATT 60
 Db |||||
 QY 641 ATTGAAAAAGATTCAAAGCGTTCCAAACAATATGATATTGAAAGATTTCGTAACATT 582
 Db |||||
 QY 61 CTACTCAGCGTATCCCATGCTTTTAGTGGATAGATATATAGATTACAAACCAATAA 120
 Db |||||
 QY 581 TTGCGCTCATAGGTTTCCGTTCTTCTAGTTGATAGATGATTAATGATCTCGAGTT 522
 Db |||||
 QY 121 AAAATTGTCGCTTATAAGAAATCACTTTTATGAAGAGCTGTTTAAACGGGCAATTCCT 180
 Db |||||
 QY 521 TCGGGCGTTGCTATTAGATAGTTAGATTAATGATTAATTTCTTCCAGGCCATTTCT 462
 Db |||||
 QY 181 AATAAGCCCAATTTTCCCGGGCTTTTGTATCGTAGAGGCGATCGCAACCGGAGGTTT 240
 Db |||||
 QY 461 GAGAGGCCAATTAATGCTGCTGTTCTTATGATTGAGGCAATGCTCAAGTTGGAGGCTTG 402
 Db |||||
 QY 241 TTAGCCTTCACATAGCTTGTGGGGGTTTGACCTGAAATCGCCAAACAAAATTCGTAT 300
 Db |||||
 QY 401 GTAATGCTGAGCCCAAGAGTGGGTGTTCTCGTG-----AAAATTTCTTC 357
 Db |||||

QY 301 TTCAATGACGATTGATAAGGTTAAATTCCGCATCCCTGTAAACCCAGGCGACAGATTAGAA 360
 Db |||||
 QY 356 TTTACAGGAATTGACAAAGTCAGATTCGGAACCTGTGGTTCAGGGGATACGCTTTGTA 297
 Db |||||
 QY 361 TACCATTAGAAAGTCTTTAAAGCATAAGGGCATGATCTGCCAAGTGGTGGCAGCGCTCAA 420
 Db |||||
 QY 296 ATGAGATGATTTCTCATTAAGAAACGCTTTGGAGTAGTTAAGATGAGGGGAAGCCTAT 237
 Db |||||
 QY 421 GTGATGCGCAAGTGGTGGTGAAGCCGAATTGAAGCCATGATTGCGAGAGAGATTAA 480
 Db |||||
 QY 236 GTTGGGGTGGAGTGGTGTATGTGACGAGAGTTACGATGGTCTTTGGGAAGAGCGAATGA 177
 Db |||||
 RESULT 8
 DR938067/c
 LOCUS
 DEFINITION EST1129606 Aquilegia cDNA library Aquilegia formosa x Aquilegia
 pubescens cDNA clone CO1PL21, mRNA sequence.
 ACCESSION DR938067
 VERSION DR938067.1 GI:71707430
 KEYWORDS EST.
 SOURCE
 ORGANISM
 Aquilegia formosa x Aquilegia pubescens
 Aquilegia formosa x Aquilegia pubescens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
 Ranunculaceae; Aquilegia.
 REFERENCE 1 (bases 1 to 841)
 AUTHORS Hodges S.A., Rensink W., Buell C.R., Borevitz J., Kramer E.,
 Nordborg M. and Tomkins J.
 TITLE Generation of ESTs from Aquilegia
 JOURNAL Unpublished (2005)
 COMMENT Other ESTs: EST1129607
 Contact: Scott Hodges
 Department of Ecology, Evolution and Marine Biology
 University of California, Santa Barbara
 Santa Barbara, CA 93106, USA
 Tel: 805 893 7813
 Fax: 805 893 4724
 Email: hodges@lifesci.ucsb.edu
 Seq primer: TTTTATTTTATTTTATTTTATTTT (where N = A, G & C).
 FEATURES
 source
 1..841
 /organism="Aquilegia formosa x Aquilegia pubescens"
 /mol_type="mRNA"
 /db_xref="taxon:338618"
 /clone="CO1PL21"
 /tissue_type="mixed shoot and floral apical meristems,
 flower buds, leaves and roots"
 /lab_host="DH10B T1 (T1 and T5 phage resistance)"
 /clone_lib="Aquilegia cDNA library"
 /note="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI;
 F2, F3, and F4 lines of Aquilegia formosa x A. pubescens
 were grown from seed in greenhouses at UC Santa Barbara.
 From these plants three sets of tissue were collected: 1)
 Small flower buds (<10 mm) and very young inflorescences
 (71 & 29% by weight respectively), 2) Medium (7-20 mm) and
 large (at or near anthesis) flower buds (65 & 35% by
 weight respectively) and 3) Shoot apical meristems. A
 fourth set of tissue was collected from plants of A.
 formosa. These plants were grown from seed in sand and at
 approximately 1 month root tissue and leaf tissue of
 various developmental stages were collected (84 & 16% by
 weight respectively). Total RNA was extracted from each
 set of tissue and pooled in the following proportions:
 1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled
 total RNA, mRNA was extracted and enriched for full-length
 messages and then normalized with proprietary methods by
 Invitrogen."
 ORIGIN
 Query Match 17.3%; Score 83.2; DB 8; Length 841;
 Best Local Similarity 50.4%; Pred. No. 4.5e-13;
 Matches 242; Conservative 0; Mismatches 223; Indels 15; Gaps 1;

COMMENT

Other_ESTs: EST1123361

Contact: Scott Hodges
Department of Ecology, Evolution and Marine Biology
University of California, Santa Barbara
Santa Barbara, CA 93106, USA
Tel: 805 893 7813
Fax: 805 893 4724
Email: hodges@lifesci.ucsb.edu

Seq primer: TTTTCTTTTCTTTTCTTTTCTTTT (where N = A, G & C).
Location/Qualifiers

FEATURES

source 1. .877

/organism="Aquilegia formosa x Aquilegia pubescens"
/mol_type="mRNA"
/db_xref="taxon:338618"
/clone="CO10876"
/tissue_type="mixed shoot and floral apical meristems,
flower buds, leaves and roots"
/lab_host="DH10B T1 (T1 and T5 phage resistance)"
/clone_lib="Aquilegia cDNA library"
/notes="Vector: pCMV SPORT6.1; Site 1: EcoRI; Site 2: NotI;
F2, F3, and F4 lines of Aquilegia formosa X A. pubescens
were grown from seed in greenhouses at UC Santa Barbara.
From these plants three sets of tissue were collected: 1)
Small flower buds (<10 mm) and very young inflorescences
(T1 & 29% by weight respectively), 2) Medium (7-20 mm) and
large (at or near anthesis) flower buds (65 & 35% by
weight respectively) and 3) Shoot apical meristems. A
fourth set of tissue was collected from plants of A.
formosa. These plants were grown from seed in sand and at
approximately 1 month root tissue and leaf tissue of
various developmental stages were collected (84 & 16% by
weight respectively). Total RNA was extracted from each
set of tissue and pooled in the following proportions:
1.5X from sets 1 & 2, 1X from sets 3 & 4. From the pooled
total RNA, mRNA was extracted and enriched for full-length
messages and then normalized with proprietary methods by
Invitrogen."

ORIGIN

Query Match 17.3%; Score 83.2; DB 8; Length 877;
Best Local Similarity 50.4%; Pred. No. 4.5e-13;
Matches 242; Conservative 0; Mismatches 223; Indels 15; Gaps 1;

QY 1 ATGGAACAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAAT 60
DB 641 ATGGAAGAAGATTCAAGCGTTTCCACAATAATGGAATTTGAAGATTCTGTAACAT 582

QY 61 CTACCTCAGCGTATCCCATGCTTTTATGAGTATGAGTATGAGTATCAAGCCAATAA 120
DB 581 TTGCCTCATAGGTTTCCGTTTCTTCTAGTTGATAGATGATGAATATGATCTCGAGTT 522

QY 121 AAAATGTGCTTATAGAAATATCACTTTTAATGAAGACGCTTTTAACGGGCATTTCCCT 180
DB 521 TCGGGGTTGCTATTAAAGAAATGTACGATTAATGATAATTTCTTCCAGGCCATTTTCCT 462

QY 181 AATAAGCCATTTTCCCGGCGTTTGTATCGTAGGGCATGCGCAACGGGAGGTTT 240
DB 461 GAGAGCCAAATTATGCTGCTGTTCTTATGATTTGAGCAATGCTCAAGTTGAGGCTTG 402

QY 241 TTAGCCTTCACATAGCTTTGTGGGGTTTGAACCTTGAATCGCCAAAACAAAATTCGTGTAT 300
DB 401 GTAATGCTGACCCAGAGTGGGTGTTCTCGTG-----AAAATTTCTTC 357

QY 301 TTCATGACAAATTGATAGTTAAATTCGGCATCCCTGTAAACCCAGCGGACAGATTAGAA 360
DB 356 TTACAGGAAATTGACAAAGTCAAGTTCCGGAACCTGTGTTGCAAGGGATACGTTGTA 297

QY 361 TACCATTAGAGTCTTAAAGCATATAGGCGATGATCTGGCAAGTGGTGCAGGGCTCAA 420
DB 296 ATGAGAAATGTTCTCATTAAGAAACGCTTTGGAGTAGTTAAGATGAGGGGAAGGCGCTAT 237

QY 421 GTGGATGGCAAGTGGTTCGCTGAAGCCGGAATTTGAAGCCATGATTTGCAGAGAGATTAA 480

DB 236 GTTGGGGTGAGTGGTATGTGACGGAGATTACGATGGCTCTTGGGAAGCGAATGA 177

RESULT 11

AJ750880
LOCUS AJ750880
DEFINITION AJ750880 G00001 Gerbera hybrid cv. 'Terra Regina' cDNA clone
G000100013G01F1, mRNA sequence.

ACCESSION

VERSION AJ750880

KEYWORDS

SOURCE AJ750880.1 GI:58321534

ORGANISM

Gerbera hybrid cv. 'Terra Regina'
Gerbera hybrid cv. 'Terra Regina'
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; campanulids; Asterales; Asteraceae; Mutisioideae;
Mutisieae; Gerbera.

REFERENCE

1 (bases 1 to 465)
Laitinen,R.A., Immanen,J., Auvinen,P., Rudd,S., Alatalo,E.R.,
Paulin,L., Ainasoja,M., Kotilainen,M., Koskela,S., Teeri,T.H. and
Elomaa,P.

TITLE

Analysis of the floral transcriptome uncovers new regulators of
organ determination and gene families related to flower organ
differentiation in Gerbera hybrida (Asteraceae)

JOURNAL

PUBMED 15781570

COMMENT

Contact: Alatalo ER
Institute of Biotechnology
Viikinkaari 9, P.O. Box 56, 00014 University of Helsinki, FINLAND.

FEATURES

source 1. .465
Location/Qualifiers
/organism="Gerbera hybrid cv. 'Terra Regina'"
/mol_type="mRNA"
/cultivar="Terra Regina"
/db_xref="taxon:226891"
/clone="G000100013G01F1"
/tissue_type="late petal"
/dev stage="5-9"
/clone_lib="G00001"

ORIGIN

Query Match 17.3%; Score 83; DB 1; Length 465;
Best Local Similarity 58.8%; Pred. No. 4.5e-13;
Matches 143; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 1 ATGGAACAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAAT 60

DB 153 ATTGAATCAGGTTCCCTCGGTTTCCACCGTTCTTGTATATCAATCACAATTCGCGACAT 212

QY 61 CTACCTCAGCGTATCCCATGCTTTTATGATGATAGAAATATATAGATTACAAGCCAATAA 120

DB 213 TTACCACACGATTTCCATTTCTTTGGTGGATAGGTTGATGATACAATCCAGGTGT 272

QY 121 AAAATGTGCTTATAGAATATCACTTTTATGAAGACGTTTAAACGGGCATTTCCCT 180

DB 273 TCAGCAGTTGGCATCAAGAATGTGACTATCAATGATACTTTTCTCTGGGCATTTCTCT 332

QY 181 AATAAGCCATTTTCCCGGCGTTTTCATCGTAGGGCATGCGCAACGGGAGGTTT 240

DB 333 GAGAGGCGGATTAATGCTGCTGTTCTTATGTCGAGCAATGGCGCAAGTTGGTGGTTG 392

QY 241 TTA 243

DB 393 GTA 395

RESULT 12

DN910931/c
LOCUS DN910931
DEFINITION Cuc6_3C12 Cucumber fruit development cDNA library Cucumis sativus
cDNA, mRNA sequence.

ACCESSION

VERSION DN910931.1 GI:62895694

[illegible]

RIKEN Genomic Sciences Center
 2-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: mseki@rc.riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES Location/Qualifiers
 source 1..598
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAFL14-24-C08"
 /tissue_type="root"
 /lab_host="DH10B"
 /clone_lib="RAFL14"
 /note="Site 1: BamHI; Site 2: SalI"

ORIGIN
 Query Match 17.2%; Score 82.4; DB 1; Length 598;
 Best Local Similarity 59.3%; Pred. No. 7.1e-13;
 Matches 140; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 1 ATCGAACAAGCCATCAAACTGCAATCTCAATTTTATAGAGCATATCTTACAATT 60
 DB 210 ATCGAACAAGGTATGAGCGTTTCCGACAGTGTGACATCAATAGATACAAGAGATT 269
 QY 61 CTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGATTACAAGCAATAA 120
 DB 270 TTGCCTCAGAGATCCCATCTCTGTTAGTGTAGATGATAGATACACAGCTGGTGA 329
 QY 121 AAAATTGCTGTTATAAGAAATACATCTTTAATGAAGAGCGTGTAAACGGGCATTTCCCT 180
 DB 330 TCTGCGGTAGCTATTAAAAAGTTACCAATTAATGACAAATTTCTTCTGGGCATTTCCCT 389
 QY 181 AATAAGCCATTTTCCGGCGGTTTGTATGCTAGAGGCGATGCGCAACGGGAGG 236
 DB 390 GAGAGACCTAATAATGCTGGAGTTCTCATGTTGAGGCAATGGCTCAAGTGGGAGG 445

RESULT 15

CNS0ALU5 774 bp mRNA linear HTC 06-FEB-2004
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
 DEFINITION GSLTSLI92P06 of Silique of strain col-0 of Arabidopsis thaliana
 (thale cress).

ACCESSION BX834110

VERSION BX834110.1 GI:42459333

KEYWORDS HTC; GSLT_cDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

AUTHORS

Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,
 Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,
 Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
 Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
 A Combined Approach to Evaluate and Improve Arabidopsis Genome
 Annotation

JOURNAL

Unpublished

REFERENCE

AUTHORS

TITLE

Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr

COMMENT

The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli
 V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
 length

FEATURES Location/Qualifiers
 source 1..774
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="GSLTSLI92P06"
 /tissue_type="Silique"
 /ecotype="Col-0"
 /plasmid="pCMVSPORT_6"
 1..774
 /gene="At5g10160"

ORIGIN

Query Match 17.2%; Score 82.4; DB 4; Length 774;
 Best Local Similarity 59.3%; Pred. No. 7.5e-13;
 Matches 140; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 1 ATGGAACAAGCCATCAAACTTGGCAATCTCAATTTTATAGAGCATATCTTACAATT 60
 DB 156 ATCGAACAAGGTATGAGCGTTTCCGACAGTGTATGACATCAATAGATACAAGAGATT 215
 QY 61 CTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGATTACAAGCAATAA 120
 DB 216 TTGCCTCAGAGATCCCATTTCTGTTAGTGTAGATGATAGATACACAGCTGGTGA 275
 QY 121 AAAATTGCTGCTTATAAGAAATACATCTTTAATGAAGACGTTGTTAAACGGGCATTTCCCT 180
 DB 276 TCTGCGGTAGCTATTAAAAACGTTACCATTAATGACAAATTTCTTCTGGGCATTTCCCT 335
 QY 181 AATAAGCCATTTTCCGGCGGTTTGTATGCTAGAGGCGATGCGCAACGGGAGG 236
 DB 336 GAGAGACCTAATAATGCTGGAGTTCTCATGTTGAGGCAATGGCTCAAGTGGGAGG 391

Search completed: January 12, 2006, 03:12:08
 Job time : 3074 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 01:14:29 ; Search time 153 Seconds
(without alignments)
5576.664 Million cell updates/sec

Title: US-10-662-126-36

Perfect score: 480

Sequence: 1 atggaacaaagccatcaaaaa.....tgattgcagagagagattaa 480

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/1/ina/1_COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5_COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/H_COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/PCUS_COMB.seq.*
- 7: /cgn2_6/prodata/1/ina/PP_COMB.seq.*
- 8: /cgn2_6/prodata/1/ina/RE_COMB.seq.*
- 9: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86.4	18.0	429	3	US-09-134-000C-1148
2	82.8	17.2	465	3	US-09-252-991A-4232
3	82.8	17.2	483	3	US-09-252-991A-4647
4	81.8	17.0	489	3	US-09-543-681A-2455
5	78.2	16.3	561	3	US-09-489-039A-5222
6	77.2	16.1	423	3	US-09-583-110-1245
7	77.2	16.1	441	3	US-09-107-433-289
8	76.6	16.0	558	3	US-09-540-236-1553
9	76.6	16.0	9829	3	US-09-596-002-37
10	75.6	15.8	423	3	US-09-196-388-1
11	75.6	15.8	483	3	US-10-089-019-17
12	75.6	15.8	19702	3	US-08-961-527-7
13	74.2	15.5	10303	3	US-09-634-238-410
14	73.4	15.3	462	3	US-09-902-540-3820
15	73.4	15.3	21010	3	US-09-902-540-1188
16	72.2	15.0	1830121	3	US-09-557-884-1
17	72.2	15.0	1830121	3	US-09-643-990A-1
18	72.2	15.0	1830121	3	US-10-158-865-1
19	69.6	14.5	461	3	US-09-339-614-1
20	69.6	14.5	501	3	US-10-089-019-9
21	69.2	14.4	471	3	US-09-134-001C-1397
22	65.4	13.6	1029	3	US-09-221-017B-988
23	62.2	13.0	492	3	US-09-328-352-3116
24	60.8	12.7	2926	3	US-09-710-279-3747

25	60.8	12.7	6309	3	US-08-956-171E-480	Sequence 480, App
26	60.8	12.7	6309	3	US-08-781-986A-480	Sequence 480, App
27	59.6	12.4	354	3	US-09-107-532A-1534	Sequence 1534, App
c 28	58.4	12.2	1230025	3	US-09-198-452A-1	Sequence 1, Appli
c 29	58.4	12.2	1230230	3	US-09-438-185A-1	Sequence 1, Appli
30	56.6	11.8	471	3	US-09-902-540-4150	Sequence 4150, Ap
c 31	56.6	11.8	18469	3	US-09-902-540-1205	Sequence 1205, Ap
32	44	9.2	591	3	US-09-543-681A-2590	Sequence 2590, Ap
33	35.6	7.4	822	3	US-09-710-279-727	Sequence 727, App
34	35.6	7.4	3215	3	US-09-710-279-3566	Sequence 3566, Ap
c 35	35.6	7.4	3993	3	US-09-710-279-3985	Sequence 3985, Ap
36	35	7.3	288031	3	US-09-949-016-14864	Sequence 14864, A
37	34.6	7.2	640681	3	US-09-790-988-1	Sequence 1, Appli
c 38	34.4	7.2	601	3	US-09-949-002-9724	Sequence 9724, Ap
c 39	34.4	7.2	36677	3	US-09-949-002-832	Sequence 832, App
c 40	34.2	7.1	7218	2	US-08-232-463-14	Sequence 14, Appl
c 41	34	7.1	412	3	US-09-270-767-3889	Sequence 3889, Ap
c 42	34	7.1	412	3	US-09-270-767-19171	Sequence 19171, A
c 43	34	7.1	83178	3	US-09-949-016-14606	Sequence 14606, A
c 44	33.8	7.0	969	3	US-09-252-991A-4308	Sequence 4308, Ap
c 45	33.6	7.0	16995	3	US-08-961-527-82	Sequence 82, Appl

ALIGNMENTS

RESULT 1

US-09-134-000C-1148
; Sequence 1148, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1148
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-1148

Query Match		18.0%;	Score 86.4;	DB 3;	Length 429;
Best Local Similarity		53.6%;	Pred. No. 5.4e-18;		
Matches		236;	Conservative	0;	Mismatches 186; Indels 18; Gaps 2;
Qy	32	AAATTTTATAGACATATCTTACAAATTTACCTCACCGCTATCCCATGCTTTTAGTGG	91		
Db	8	AAATTAACAATTTACAGAAATTTCAAGAAATTTTACCATCGCTATCCCTTTTATTATTAG	67		
Qy	92	ATAGAATTTATAGAGTTACAAAGCAATAAAAAATTTGCTTTATAAGAATATCACTTTTA	151		
Db	68	ATAGTGTGTGAAGAGTCAATTCGCGGGAACGCGTTGTAGCGAAAGAAATGTAACGGTTA	127		
Qy	152	ATGAAGACGTGTTTAAACGGGCAATTTCCCTAATAAGCCATTTCCCGGCGGTTTGTATCG	211		
Db	128	ATGAGCAAGTTTTTCAAGGTCATTTTCTCGGAATCTGTTTTCGCGAGAGTTTTTAATTA	187		
Qy	212	TAGAGGCATGGCCAAACGGAGGTTTTTAGCTTCACTAGCTTGTGGGGGTTTGACC	271		
Db	188	TTGAATCGTTGGCTCAAGCAGGCGCGCTGGCATTTCTCAA-----TGC	232		
Qy	272	CTGAATTCGCAAAACAAAATCTGTATTTCATGACGATTGATAAGTGTAAATTCGCA	331		
Db	233	CTGAATTT---CAAGGGAAACCGGCATCTTTGGTGAATTAGTAAGCAAAATTTAGAC	289		
Qy	332	TCCTGTAAACCCAGCGCAGACGATTAGAAATCCATTAGAGTCTTAAAGCATAGGGCA	391		

Db 290 AAAAGGTAACACCGGACACATTAATTTTAGAGGTTGAACCTTTTAAAGATCCGGCTT 349
QY 392 TGATCTGCAAGTGGGTGGCAGCGCTCAAGTGGATGGCAAGTGGTGGCTGAAGCCGAAT 451
Db 350 CTGCTGGAATGGTAAAGTGTGGCAAAAGTTAATGTAAGAAAAAGTTGCCGAGGCTGAAT 409
QY 452 TGAAGCCATGATTCGAG 471
Db 410 TAACTTTATGATGGATAG 429

RESULT 2
US-09-252-991A-4232/c
; Sequence 4232, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4232
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4232

Query Match 17.2%; Score 82.8; DB 3; Length 465;
Best Local Similarity 54.2%; Pred. No. 8.4e-17;
Matches 218; Conservative 0; Mismatches 172; Indels 12; Gaps 2;

QY 56 AAATCTACTCACCGCTATCCCATGCTTTTAGTGATAGAAATTAAGAGTTTACAAGCCA 115
Db 410 AATACCTGCTCATCGTATCCTTCTCTGCTGGTGGATCGGTGGTAGAGTGGACATCG 351
QY 116 A---TAAAAAATGTGCTTTAAGAAATATCACTTTTATGAAGACGTGTTTAACGGGC 172
Db 350 AGGCAAGCGCATTCGCGCTACAAGATGTCAATGATCAATGAGCGGTCTTCAACGGAC 291
QY 173 ATTTCCCTAATAGCCATTTTCCGGGGCTTTTGATGCTAGAGGCGATGGCCAAACGG 232
Db 290 ATTTCCCGAGCACCGCATCATGCGGGGGTGTGCTGATCATCGAGGCGATGGCCGCGG 231
QY 233 GAGGGTTTTAGCTTCACTAGCTTTGCGGGGTTTGACCTGAAATCGCAAAACAAAA 292
Db 230 CCGGTATCTCGTTTCA-----AGATGCTCGAGTGAAGCCGCGACGCCACCC 180
QY 293 TCGTGTAATTCATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 352
Db 179 TTTACTACTTCTGCTGGTTCGGAAGTGTGGTCTTCCGCGAGCGGTCTTCCGCGGAC 120
QY 353 GATTGAATACCATTTAAGAGTCTTAAAGCAATGAAGGCGATGATCTGGCAAGTGGTGGCA 412
Db 119 AACTGCAACTGCACGCAAGTTCATCAGCGTGAAGCGCAGCATCTGGAATTCGACTGCC 60
QY 413 CGGCTCAAGTGGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 454
Db 59 ACGCTACCGTGCAGCAAGCGGTATGCTCGGCTGAAATCA 18

RESULT 3
US-09-252-991A-4647
; Sequence 4647, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4647
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4647

Query Match 17.2%; Score 82.8; DB 3; Length 483;
Best Local Similarity 54.2%; Pred. No. 8.6e-17;
Matches 218; Conservative 0; Mismatches 172; Indels 12; Gaps 2;

QY 56 AAATCTACTCACCGCTATCCCATGCTTTTAGTGGATAGAATTAAGATTACAGGCCA 115
Db 68 AATACCTGCTCATCGCTATCCTTTCCTGCTGGTGGATCGGTGGTAGAGCTGGACATCG 127
QY 116 A---TAAAAAATGTGCTTTAAGAAATATCACTTTTATGAAGACGTGTTTAACGGGC 172
Db 128 AGGCAAGCGCATTCGCGCTACAAGAAATGTCAATGATCAATGAGCCGTTCTTCAACGGAC 187
QY 173 ATTTCCCTAATAGCCATTTTCCGGGGCTTTTGATGCTAGAGGCGATGGCCAAACGG 232
Db 188 ATTTCCCGAGCACCGCATCATGCGGGGGTGTGATCATCGAGGCGATGGCCAGGCGG 247
QY 233 GAGGGTTTTAGCTTCACTAGCTTTGCGGGGTTTGACCTGAAATCGCAAAACAAAA 292
Db 248 CCGGTATCTCGTTTCA-----AGATGCTCGAGTGAAGCCGCGACGCCACCC 298
QY 293 TCGTGTAATTCATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 352
Db 299 TTTACTACTTCTGCTGGTTCGGAAGTGTGGTCTTCCGCGAGCGGTCTTCCGCGGAC 358
QY 353 GATTGAATACCATTTAAGAGTCTTAAAGCAATGAAGGCGATGATCTGCAAGTGGTGGCA 412
Db 359 AACTGCAACTGCAGCCCAAGTTTATCAGCGTGAAGCGCAGCATCTGGAATTCGACTGCC 418
QY 413 CGGCTCAAGTGGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 454
Db 419 ACGTACCGTGCAGCAAGCGGTATGCTCGGCTGAAATCA 460

RESULT 4
US-09-543-681A-2455
; Sequence 2455, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2455
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-2455

Query Match 17.0%; Score 81.8; DB 3; Length 489;
Best Local Similarity 52.0%; Pred. No. 1.8e-16;
Matches 216; Conservative 0; Mismatches 187; Indels 12; Gaps 1;

QY 40 ATAGAGCATATCTTACAAATTTCTACCTCACCGCTATCCCATGCTTTTAGTGGATGAAT 99

Db 61 ATCGAAGAAATTTTAGATTCTTACCTCCACCGTTACCAATTTTATTTGTTGACCGTGC 120
Qy 100 ATAGAGTTAAGCAATATAAATAATTTGCTGTTATAGAAATATCACTTTTAAATGAAGAC 159
Db 121 CTGTGATTTTGAAGAAGGAAATTTTAAAGAGCAGTAAAAAATGTATCTTTTAAATGAGCCT 180
Qy 160 GTGTTTAAACGGGATTTCCCTAATAGCCCATTTTCCGGGGCGTTTGTATCGTAGAGGC 219
Db 181 TTCTTTAAAGGTCACTTCCAGCAAACTTAATTTCCCTGGTGTACTGATCTCTGGAAGCA 240
Qy 220 ATGCGCAAAACGGGAGGTTTGTAGCTTCACTAGCTTGTGGGGGTTTGAACCTTGAATC 279
Db 241 ATGCGCGAAGCAGAGTATTTTAGCATTCAAAGCGTAGAAAACTGGAAACCG----- 295
Qy 280 GCCAAACAAAAATCGTGATTTTCATGCAATGATAAGGTTAAATTCGCAATCCCTGTA 339
Db 296 -----GTGAATTTTACTATTTTGTCTGCTATTGACGGTCTCGCTTTAAAGGCTCTGTC 348
Qy 340 ACCCGAGGCGACAGATTAGAATACCATTTAGAGTCTTAAAGCATTAAGGCGATGATCTGG 399
Db 349 TTACCTGGAGACCAAAATGTTCTTGAGGTTGAATTCATTAAGAAACCGCGTGGCGTTGCA 408
Qy 400 CAAAGTGGTGGCAGCGCTCAAGTGGATGCAAGTGTGCTGCTGAAGCGCAATTGA 454
Db 409 CGCTTTAAAGGCGCTGCTAAGTTGATGAGAAATTTGCCGCGAAGCTGAGATGA 463

RESULT 5
US-09-489-039A-5222
; Sequence 5222, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5222
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5222

Query Match 16.3%; Score 78.2; DB 3; Length 561;
Best Local Similarity 58.3%; Pred. No. 3e-15;
Matches 137; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 40 ATAGAGCATATCTTACAAATTTTACTCACCCTATCCCATGCTTTTAGTGGATAGATT 99
Db 133 ATGAAAGAAATTTGGAACCTTCTGCTCAACCGTTACCGCTTCTGCTGGTAGACCGGTG 192
Qy 100 ATAGAGTTACAAGCCAAATAAATAATTTGCTGTTATAAGCAATATCACTTTTATGAAGAC 159
Db 193 CTGGATTTTGAAGAGGTGCTTTCTGCGCGCAGTAAATAATGTTCTGTAAACGAGCGG 252
Qy 160 GTGTTTAAACGGGCAATTCCTAATAAGCCATTTTCCGGGGCGTTTGTATCGTAGAGGC 219
Db 253 TTTTTCAGGGGCACTTCTCTGTGAAGCGATTTTGCAGGGCTCTGATCTCTGGAAGCA 312
Qy 220 ATGCGCAAAACGGGAGGTTTGTAGCTTCACTAGCTTGTGGGGGTTTGACCCCTG 274
Db 313 ATGGCGCAAGCAGCCGGAATTTGCGGTTTCAAAAGCGTCCGAAAACTGGAACACAG 367

RESULT 6
US-09-583-110-1245
; Sequence 1245, Application US/09583110
; Patent No. 6699703

; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 1245
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-1245

Query Match 16.1%; Score 77.2; DB 3; Length 423;
Best Local Similarity 55.5%; Pred. No. 5.4e-15;
Matches 223; Conservative 0; Mismatches 158; Indels 21; Gaps 3;

Qy 39 TATAGACATATCTTACAAATTTCTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAT 98
Db 9 TATTCAGGAATCAAGAAGCCCTTCCCAACCGTTATCTCTATGCTTCTAGTGGACCGTGT 68
Qy 99 TATAGAGTTACAGCCCAATAAATAATTTGCGCTTATAAGAAATATCACTTTTAAATGAAGA 158
Db 69 CTGGAGGT---GAGCGAGGATACCATTTGCTTATCAAAAATGTGACCATCAACGAGCC 125
Qy 159 CGTGTTTTAAACGGGCAATTTCCCTAATAAGCCCAATTTTCCGGGGCGTTTGTATCGTAGAGGG 218
Db 126 TTTCTTTTAAACGGGCAATTTCTCAATACCCAGTTATGCCAGGTGTTCTGATATGAAGAC 185
Qy 219 CATGGCGCAAAACGGGAGGTTTGTAGCTTCACTAGCTTGTGGGGGTTTGAACCTGAAAT 278
Db 186 CTTGGCGCAAACTGCCGTGTTGGAGTTATCAA-----AACCTGAAA- 229
Qy 279 CGCCAAACAAATAATCGTGATTTTCATGACGATGATAAGTTAAATTCGCAATCCCTGT 338
Db 230 --ATAAGGAAATCTGCTTTTACGCTGATGATGAAGTTAAAGTTCAAGAAGCAAGT 287
Qy 339 AACCCAGGCGCACAGATTAGATATACCATTTAGAAAGTCTTAAAGCATTAAGGCGCATGCTG 398
Db 288 TGTACCAAGCGCACCAATTTGTTATGACAGCGACTTTTGTAAACGTCGTGGCACCATAGC 347
Qy 399 GCAAGTGGGTGGCAGCGCTCAAGTGGATGGCAAGTGGTGGC 440
Db 348 TGTGTTGAAGCAAGGCTGAAGTGGATGGCAAGCTTGCAGC 389

RESULT 7
US-09-107-433-289
; Sequence 289, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: <Unknown>

OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 289:
SEQUENCE CHARACTERISTICS:
LENGTH: 441 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...441
SEQUENCE DESCRIPTION: SEQ ID NO: 289:
US-09-107-433-289

Query Match 16.1%; Score 77.2; DB 3; Length 441;
Best Local Similarity 55.5%; Pred. No. 5.5e-15;
Matches 223; Conservative 0; Mismatches 158; Indels 21; Gaps 3;

QY 39 TATAGGATATCTTACAAATTTCTACCTCACCGCTATCCCGCTTTTAGTGATAGAT 98
DB 27 TATTCAGGAATCAAGAAGCCCTTCCCGCCGTTATCTCTATGACCGTGT 86

QY 99 TATAGAGTTTACAGGCAATATAAATAATTCGCTTATAAGATATACATTTTAAATGAAGA 158
DB 87 CTGGAGGT--GAGCGAGGATACCATTTGCTATCAAAATGTGACCATCAACGAGCC 143

QY 159 CGTGTTAACGGGCAATTCCTTAATAAGCCCAATTTCCCGGGCGTTTGTATCGTAGAGGG 218
DB 144 TTTCTTTAAGCGCCACTTTCTCTCAATACCCAGTTATGCCAGGTGTCTGTATATGAAGC 203

QY 219 CATGGGCAACGGGAGGGTTTTCAGCTTCACTAGCTTTGCGGGGTTTGACCCCTGAAAT 278
DB 204 CTGGCGCAACTGCGGGTGTGTGAGTTATCAA-----AACCTGAAA- 247

QY 279 CGCCAAAACAAAATCGTGTATTTATGACGATTAAGGTTAAATTCGCAATCCCTGT 338
DB 248 --ATAAGAAATCTGTCTTTTACCTGGTATGATAGGTTAAGTTCAAGAGCAAGT 305

QY 339 AACCCAGCGCAGATTAAGATACCATTTAGAAAGTCTTAAAGCATAAAGGCGATGATCTG 398
DB 306 TGTACAGCGCAACCAATTTGTTATGACAGGCACTTTTGTAAACGCTGTGGCCACCATAGC 365

QY 399 GCAGTGGGTGGCAGCGCTCAAGTGAATGGCAAGTGGTGC 440
DB 366 TGTGGTGAAGCAAGGCTGAAGTGGATGGCAAGCTTGCAGC 407

RESULT 8
US-09-540-236-1553
; Sequence 1553, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRH
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 1553
; LENGTH: 558
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-1553

Query Match 16.0%; Score 76.6; DB 3; Length 558;
Best Local Similarity 51.9%; Pred. No. 9.8e-15;
Matches 203; Conservative 0; Mismatches 179; Indels 9; Gaps 1;

QY 44 AGCATATCTTACAAATTTCTACCTCACCGCTATCCCGCTTTTAGTGATAGATATATAG 103
DB 122 ACCAAATTAAGCATTAATTTGCCGCATCGCTATCCATTCATGCTGATGACCGAGCGG 181

QY 104 AGTTACAAGCCATAAAAAAATTTGCTGCTTATAAGAAATATCAGCTTTTAAATGAAGACGTGT 163
DB 182 CTGTAAACCCCAATGAGTGATCAGGATATAAAAAATATCAGTATTAAATGAAGATTAT 241

QY 164 TTAACGGGCAATTTCCCTTAATAAGCCCAATTTTCCCGGGCGTTTTCATCGTAGAGGCGATGG 223
DB 242 TTAACGGTCAATTTTCCAGATAATCCCAATCATGCGAGCGTACTACAAGTTTGAAGCAATGG 301

QY 224 CGCAACGGGGAGGGTTTTCAGCTTACATGAGCTTTGCGGGGTTTGACCCCTGAAATCGCCA 283
DB 302 CACAGTATCGGGTATTTTAGGTTTTCATGAGTGAGGGCAGACCGCTGATG----- 352

QY 284 AAACAAAAATCGTGATTTTCATGACGATTCATAGAGTTAAATTCGCGATCCCTGTAAACCC 343
DB 353 ATGTTATCTGTACTTGTGTCAGGTGTGATAAAGTTTCGTTTAAAAAGTCTGTACCT 412

QY 344 CAGGCGACAGATTAGAAATACCATTTAGAACTTTAAAGCATAAAGGCGATGATCTGGGCAAG 403
DB 413 CAGGCGATCAGCTGGTGTATTTCGCTCAAGGTTTTCATGATGAATAAACGAGATATTTATAAAT 472

QY 404 TGGGTGCGCAGCGCTCAAGTGATGGCAAGT 434
DB 473 TTGAGTGTACCGCCCATGTTGATGGTCAATT 503

RESULT 9
US-09-596-002-37
; Sequence 37, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 37
; LENGTH: 95629
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 78467
; OTHER INFORMATION: a or g or c or t, unknown, or other
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte template ID No. 6632636 37
; PUBLICATION INFORMATION:
US-09-596-002-37

Query Match 16.0%; Score 76.6; DB 3; Length 99629;
Best Local Similarity 51.9%; Pred. No. 1.4e-13;
Matches 203; Conservative 0; Mismatches 179; Indels 9; Gaps 1;

QY 44 AGCATATCTTACAAATCTACCTCACCGCTATCCATGCTTTTAGTGAGATAGATATAG 103
DB 69789 ACCAAATTAAGCATTAATTTCCCGATGCTATCAATCACTGATGACCGAGTACCG 69848

QY 104 AGTTCAAGCAATAAAAAATTTGCTTTTATAGAAATATCACTTTTAAATGAAGACGTGT 163
DB 69849 CTTGTAAACCAATGATGATGATACAGGATATAAAAAATATCACTTTTAAATGAAGATAT 69908

QY 164 TTAACGGCAATTTCCCTAATAAGCCATTTTCCGGCGCTTTTGAATCGTAGAGGCGATGG 223
DB 69909 TTAACGGCTCATTTTCCAGATAATCCAAATCATGCGCGCTACTACAAAGTTGAAGCAATGG 69968

QY 224 CGCAACCGGAGGCTTTTAGCTTCCATAGCTTGTGGGGTTTGACCTCGAAATCGCCA 283
DB 69969 CACAGCTATCGGTATTTTAGTGTTCATGCTGAGGCGAGCGCTGATG----- 70019

QY 284 AAACAAAAATCGTGTATTTTCATGACGATTTGAAGTTAAATTCGCGATCCCTGTAAACCC 343
DB 70020 ATGTTTATCTGTACTTGTGTCAGGTGTCGATAAGTTCGTTTAAAAAAGTCTTACCT 70079

QY 344 CAGGCGACAGATTAGAAATACCATTTTGAAGTCTTAAAGCATAAGGCGATGATCTGGCAAG 403
DB 70080 CAGGCGATCAGCTGCTTATTCGCTCAAAGGTTTGTATGAATAAACGAGATATTTATAAAT 70139

QY 404 TGGGTGCGACGCTCAAGTGGATGGCAAGT 434
DB 70140 TTGAGTGTACCGCCCATGTTGATGGTCAATT 70170

RESULT 10
US-09-196-388-1
; Sequence 1, Application US/09196388
; Patent No. 6277595
; GENERAL INFORMATION:
; APPLICANT: Warren, Patrick
; APPLICANT: Konstantinidis, Alex
; APPLICANT: Russell, Robert
; TITLE OF INVENTION: fab2
; FILE REFERENCE: GM10182
; CURRENT APPLICATION NUMBER: US/09/196,388
; CURRENT FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(420)
US-09-196-388-1

Query Match 15.8%; Score 75.6; DB 3; Length 423;
Best Local Similarity 55.2%; Pred. No. 1.8e-14;
Matches 222; Conservative 0; Mismatches 159; Indels 21; Gaps 3;

QY 39 TATAGAGCATATCTTACAAATCTACCTCACCGCTATCCCATGCTTTTAGTGATAGAT 98
DB 9 TATTCAGGAATCAAGAAGCTCTTCCCAACGTTATCTATCTATGCTTCTAGTGGACCGTGT 68

QY 99 TATAGAGTTACAGCCCAATAAAAAATTTGCTTTTAAAGAAATATCACTTTTAAATGAAGA 158
DB 69 CTTTGA---AGTGAGGAGGATACCATTTGCTATCAAAAATGTGACCATCAACGAGCC 125

QY 159 CGTGTTTAAGGGCATTTCCCTAATAAGCCCATTTTCCGGCGCTTTTGTATCGTAGAGGG 218
DB 126 TTTCTTTAAGCGCCATTTCTCTCAATACCCAGTTATGCCAGTCTGTGTGATTATGGAAGC 218

QY 219 CATGGCGAAACGGAGGGTTTTTAGCTTCACTAGCTTGTGGGGTTTTGACCTGAAAT 278
DB 186 CTTGGCGAAACTGCCGTGTGTTGAGTTATCAA-----AACCTGAAA- 229

QY 279 CGCCAAAACAAAAATCGTGTATTTTCATGACGATTTGAAGTTAAATTCGCGATCCCTGT 338
DB 230 --ATAAGGAAAACTGGTCTTTTACGCTGTTATGCAAGGTTAAGTTTCAAGAAGCAAGT 287

QY 339 AACCCAGGCGACAGATTAGAAATACCATTTTGAAGTCTTAAAGCATAAGGGCATGATCTG 398
DB 288 TGTACCAAGCGACCAATTTGGTTATGACAGCGACTTTTGTAAAAACGTCGTGGCACCATAGC 347

QY 399 GCAAGTGGTGGCAGCGCTCAAGTGGATGGCAAAAGTGTGCGC 440
DB 348 TGTGTTGAAGCAAGGCTGAAGTGGATGGCAAGCTTGCAGC 389

RESULT 11
US-10-089-019-17
; Sequence 17, Application US/10089019
; Patent No. 6951729
; GENERAL INFORMATION:
; APPLICANT: DEWOLF, WALTER E. JR
; APPLICANT: KALLENDER, HOWARD
; APPLICANT: LONSDALE, JOHN T.
; TITLE OF INVENTION: METHODS FOR MAKING AND USING FATTY ACID
; FILE REFERENCE: GM50068
; CURRENT APPLICATION NUMBER: US/10/089,019
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: PCT/US00/29451
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/161,775
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-089-019-17

Query Match 15.8%; Score 75.6; DB 3; Length 483;
Best Local Similarity 55.2%; Pred. No. 1.9e-14;
Matches 222; Conservative 0; Mismatches 159; Indels 21; Gaps 3;

QY 39 TATAGAGCATATCTTACAAATCTACCTCACCGCTATCCCATGCTTTTAGTGATAGAT 98
DB 69 TATTCAGGAATCAAGAAGCTCTTCCCAACGTTATCTATGCTTCTAGTGGACCGTGT 128

QY 99 TATAGAGTTACAGCCCAATAAAAAATTTGCTTTTAAAGAAATATCACTTTTAAATGAAGA 158
DB 129 CTTTGA---AGTGAGCGAGGATACCATTTGCTATCAAAAATGTGACCATCAACGAGCC 185

QY 159 CGTGTTTTACCGGCAATTTCCCTAATAAGCCCATTTTCCGGCGCTTTTGTATCGTAGAGGG 218
DB 186 TTTCTTTTACCGCCACTTCTCTCAATACCCAGTTATGCCAGTCTGTGTGATTATGGAAGC 245

QY 219 CATGGCGAAACGGAGGGTTTTTAGCTTCACTAGCTTGTGGGGTTTTGACCTGAAAT 278
DB 246 CTTGGCGCAACTGCCGCTGTGTTGGAGTTATCAA-----AACCTGAAA- 289

QY 279 CGCCAAAACAAAAATCGTGTATTTTCATGACGATTTGAAGTTAAATTCGCGATCCCTGT 338
DB 290 --ATAAGGAAAACTGGTCTTTTACGCTGTTATGCAAGGTTAAGTTTCAAGAAGCAAGT 347

QY 339 AACCCAGGCGACAGATTAGAAATACCATTTTGAAGTCTTAAAGCATAAGGGCATGATCTG 398
DB 348 TGTACCAAGCGACCAATTTGTTATGACAGCGACTTTTGTAAAAACGTCGTGGCACCATAGC 407

QY 399 GCAAGTGGTGGCAGCGCTCAAGTGGATGGCAAAAGTGTGCGC 440
DB 408 TGTGTTGAAGCAAGGCTGAAGTGGATGGCAAGCTTGCAGC 449

US-09-902-540-3820

Query Match 15.3%; Score 73.4; DB 3; Length 462;
 Best Local Similarity 50.1%; Pred. No. 9.9e-14;
 Matches 219; Conservative 0; Mismatches 206; Indels 12; Gaps 1;
 QY 40 ATAGAGCATATCTTACAAATTCCTACCGCTATCCCATGCTTTTATGATAGATTA 99
 DB 7 ATCGGAGAGATTCATCTGCTGCGCATCGGTACCGCTTCTGCTGGGACCGGGTG 66
 QY 100 ATAGAGTTTCAAGCCCAATAAAAAATTTGCTTTATAGAAATATCACTTTTATGAAGAC 159
 DB 67 GTGGAGATCATCCCGGCCAGAGCTGACGGCTTACAGAACTCACAGAGGCC 126
 QY 160 GTGTTTAAAGGGCATTTCCCTAATAAGCCCATTTTCCCGGGGCTTTTATGATGAGGGC 219
 DB 127 TTCTTCAAGGGGCATTTCCCGGTACCCCGGTGATCCCGGTGCTCATCTCTGAGGGC 186
 QY 220 ATGGCGCAAGGGAGGGTTTATGCTTTCACTAGCTTTGAGGGGTTTGAACCTGAAATC 279
 DB 187 CTGGCGCAGGCGAGCCCATCTCGCTTACAGAGCGGAGAACATGGAT----- 234
 QY 280 GCCAAACAAAATCTGTATTTATGACGATTTGATAAGGTTAAATTCGCGATCCCTGTA 339
 DB 235 CGTCCGCGAAGCTCACGTACCTGATGGCGGTGAGCGCGCGCTTCCGCAAGCCGGTG 294
 QY 340 ACCCGAGGCGACAGATTAGAATACCAATTTAGAGTCTTAAAGCATAGGGCGCATGCTGG 399
 DB 295 TTGCGGGGGATCGCCTCCAGTTGGAGATCGAGGTGTCGTCACAAAGGGCGCTGCTGG 354
 QY 400 CAAAGTGGTGGACAGGCTCAAGTGGATGCGAAGTGGTGGTGGTGGTGGTGGTGGTGG 459
 DB 355 AAGACGAAGGGCTGGCGCATGCGCGCGCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 414
 QY 460 ATGATTGCAGAGAGAGA 476
 DB 415 ACCGTCTGGACAAGGA 431

RESULT 15
 US-09-902-540-1188
 ; Sequence 1188, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540
 ; PRIORITY FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: 60/217,883
 ; PRIORITY FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 1188
 ; LENGTH: 21010
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(21010)
 ; OTHER INFORMATION: unsure at all n locations
 US-09-902-540-1188

Query Match 15.3%; Score 73.4; DB 3; Length 21010;
 Best Local Similarity 50.1%; Pred. No. 7e-13;
 Matches 219; Conservative 0; Mismatches 206; Indels 12; Gaps 1;
 QY 40 ATAGAGCATATCTTACAAATTCCTACCGCTATCCCATGCTTTTATGATAGATTA 99
 DB 12878 ATCGGAGAGATTCATCTGCTGCGCATCGGTACCGCTTCTGCTGGTGGACCGGGTG 12937

QY 100 ATAGAGTTTCAAGCCCAATAAAAAATTTGCTTTATAGAAATATCACTTTTATGAAGAC 159
 DB 12938 GTGGAGATCATCCCGGCCAGAGCTGACGGCTTACAGAACTCACAGAGGCC 12997
 QY 160 GTGTTTAAAGGGCATTTCCCTAATAAGCCCATTTTCCCGGGGCTTTTATGATGAGGGC 219
 DB 12998 TTCTTCAAGGGGCATTTCCCGGTACCCCGGTGATCCCGGTGCTCATCTCTGAGGG 13057
 QY 220 ATGGCGCAAGGGAGGGTTTATGCTTTCACTAGCTTTGAGGGGTTTGAACCTGAAATC 279
 DB 13058 CTGGCGCAGGCGAGCGGCATCTCGCTTACAGAGCGGAGAACATGGAT----- 13105
 QY 280 GCCAAACAAAATCTGTATTTATGACGATTTGATAAGGTTAAATTCGCGATCCCTGTA 339
 DB 13106 CCGTCCGCGAAGCTCACGTACCTGATGGCGGTGAGCGCGCGCTTCCGCAAGCCGGTG 13165
 QY 340 ACCCGAGGCGACAGATTAGAATACCAATTTAGAGTCTTAAAGCATAGGGCGCATGATCTGG 399
 DB 13166 TTGCGGGGGATCGCCTCCAGTTGGAGATCGAGGTGGTGGTGGTGGTGGTGGTGG 13225
 QY 400 CAAAGTGGTGGCACGGCTCAAGTGGATGGCAAGTGGTGGTGGTGGTGGTGGTGGTGG 459
 DB 13226 AAGACGAAGGGCTGGCGACGGTGGCGCGCGGTGGCGCGCGGTGGCGCGAGGCGAGTTCCTGGCA 13285
 QY 460 ATGATTGCAGAGAGAGA 476
 DB 13286 ACCGTCTGGACAAGGA 13302

Search completed: January 12, 2006, 03:14:46
 Job time : 156 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 01:21:48 ; Search time 658 Seconds
(without alignments)
6032.373 Million cell updates/sec

Title: US-10-662-126-36

Perfect score: 480

Sequence: 1 atgggaacaaagcacaataa.....tgattgcagagagattaa 480

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.Main.*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	480	100.0	480	7	US-10-282-122A-22798
2	480	100.0	480	9	US-10-662-126-36
3	454.4	94.7	480	3	US-09-815-242-7504
4	454.4	94.7	480	7	US-10-335-977-2213
5	145.4	30.3	180	7	US-10-335-977-2212
6	134.4	28.0	441	7	US-10-282-122A-18085
7	99.2	20.7	459	6	US-10-369-493-3520
8	95	19.8	432	7	US-10-282-122A-16557
9	93.4	19.5	417	7	US-10-282-122A-21741
10	92.6	19.3	462	7	US-10-282-122A-41232
11	88.6	18.5	435	7	US-10-282-122A-24354
12	88.6	18.5	1755	7	US-10-398-221-1922
13	88.6	18.5	1927	7	US-10-398-221-3523
14	88.2	18.4	459	7	US-10-282-122A-31282
15	87.4	18.2	453	7	US-10-282-122A-14857
16	87	18.1	684707	7	US-10-398-221-9
17	87	18.1	3011208	7	US-10-398-221-2058
18	86.4	18.0	426	3	US-09-815-242-6826
19	85.4	18.0	6021	3	US-09-070-927A-458
20	86.2	18.0	414	6	US-10-369-493-35390
21	86.2	18.0	414	6	US-10-369-493-37892
22	86.2	18.0	414	6	US-10-369-493-38330
23	86.2	18.0	414	6	US-10-369-493-38792

24	86.2	18.0	426	7	US-10-282-122A-15658	Sequence 15658, A
25	85	17.7	423	7	US-10-282-122A-21199	Sequence 21199, A
26	85	17.7	471	6	US-10-369-493-23894	Sequence 23894, A
27	85	17.7	518	6	US-10-369-493-44892	Sequence 44892, A
28	83.6	17.4	420	6	US-10-369-493-44343	Sequence 44343, A
29	82.8	17.2	438	7	US-10-282-122A-33603	Sequence 33603, A
30	82.8	17.2	441	3	US-09-815-242-7839	Sequence 7839, AP
31	82.8	17.2	441	7	US-10-282-122A-30332	Sequence 30332, A
32	82.8	17.2	441	9	US-10-965-006-1	Sequence 1, Appli
33	82.8	17.2	441	9	US-10-965-006-3	Sequence 3, Appli
34	82	17.1	435	6	US-10-369-493-33616	Sequence 33616, A
35	81.8	17.0	435	6	US-10-369-493-23694	Sequence 23694, A
36	81.6	17.0	441	6	US-10-369-493-41684	Sequence 41684, A
37	81.4	17.0	399	6	US-10-369-493-47014	Sequence 47014, A
38	80.2	16.7	518	7	US-10-282-122A-32389	Sequence 32389, A
39	78.8	16.4	423	7	US-10-282-122A-37625	Sequence 37625, A
40	78.8	16.4	432	7	US-10-282-122A-9293	Sequence 9293, AP
41	78.8	16.4	3274	3	US-09-070-927A-350	Sequence 350, APD
42	78.4	16.3	432	7	US-10-282-122A-6335	Sequence 6335, AP
43	78.4	16.3	435	3	US-09-815-242-6420	Sequence 6420, AP
44	78.2	16.3	453	7	US-10-282-122A-23419	Sequence 23419, A
45	78	16.2	435	7	US-10-282-122A-25130	Sequence 25130, A

ALIGNMENTS

RESULT 1

US-10-282-122A-22798
; Sequence 22798, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22798
; LENGTH: 480

```
; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-10-282-122A-22798

Query Match      100.0%; Score 480; DB 7; Length 480;
Best Local Similarity 100.0%; Pred. No. 3.6e-131;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
   |||||
Db 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
   |||||

QY 61 CTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTCAGGCAATAA 120
   |||||
Db 61 CTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTCAGGCAATAA 120
   |||||

QY 121 AAAATTGTCGCTTATAGAATATCACTTTTAAATGAAGACGCTTTTAAACGGGCATTTCCCT 180
   |||||
Db 121 AAAATTGTCGCTTATAGAATATCACTTTTAAATGAAGACGCTTTTAAACGGGCATTTCCCT 180
   |||||

QY 181 AATAAGCCCATTTTCCCGGCGTTTGTGATCGTAGAGGCGATGCGCAAAACGGGAGGTTT 240
   |||||
Db 181 AATAAGCCCATTTTCCCGGCGTTTGTGATCGTAGAGGCGATGCGCAAAACGGGAGGTTT 240
   |||||

QY 241 TTAGCCCTTCACTAGCTTGTGGGGTTTGACCTCGAATCGCCAAAACAAAAATCGTGTAT 300
   |||||
Db 241 TTAGCCCTTCACTAGCTTGTGGGGTTTGACCTCGAATCGCCAAAACAAAAATCGTGTAT 300
   |||||

QY 301 TTCATGACGATGATAGGTTAAATTCGCGATCCCTGTAACCCAGCGACAGATTAGAA 360
   |||||
Db 301 TTCATGACGATGATAGGTTAAATTCGCGATCCCTGTAACCCAGCGACAGATTAGAA 360
   |||||

QY 361 TACCATTTAGAAGTCTTAAAGCATAGGGCATGATCTGGCAAGTGGGTGGCAGGCTCAA 420
   |||||
Db 361 TACCATTTAGAAGTCTTAAAGCATAGGGCATGATCTGGCAAGTGGGTGGCAGGCTCAA 420
   |||||

QY 421 GTGGATGGCAAGTGTGCTGAAGCGGAATTGAAGCCCATGATTGACAGAGAGATTAA 480
   |||||
Db 421 GTGGATGGCAAGTGTGCTGAAGCGGAATTGAAGCCCATGATTGACAGAGAGATTAA 480
   |||||

RESULT 2
US-10-662-126-36
; Sequence 36, Application US/10662126
; Publication No. US20050063987A1
; GENERAL INFORMATION:
; APPLICANT: Knapp, Bernhard
; APPLICANT: Hundt, Brika
; APPLICANT: Schmidt, Karl-Heinz
; TITLE OF INVENTION: Proteins, In Particular Membrane Proteins, of
; FILE REFERENCE: Helicobacter Pylori, Their Preparation and Use
; CURRENT APPLICATION NUMBER: US/10/662,126
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: 09/230,158
; PRIOR FILING DATE: 1999-01-19
; PRIOR APPLICATION NUMBER: PCT/IB97/00981
; PRIOR FILING DATE: 1997-07-25
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 36
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(480)
; OTHER INFORMATION: bp mature protein; 17 kD protein from Helicobacter
; OTHER INFORMATION: Pylori
US-10-662-126-36

Query Match      100.0%; Score 480; DB 9; Length 480;
Best Local Similarity 100.0%; Pred. No. 3.6e-131;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
   |||||
Db 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
   |||||

QY 61 CTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTCAGGCAATAA 120
   |||||
Db 61 CTACCTCACCGCTATCCCATGCTTTTAGTGGATAGAAATTATAGAGTTCAGGCAATAA 120
   |||||

QY 121 AAAATTGTCGCTTATAGAATATCACTTTTAAATGAAGACGCTTTTAAACGGGCATTTCCCT 180
   |||||
Db 121 AAAATTGTCGCTTATAGAATATCACTTTTAAATGAAGACGCTTTTAAACGGGCATTTCCCT 180
   |||||

QY 181 AATAAGCCCATTTTCCCGGCGTTTGTGATCGTAGAGGCGATGCGCAAAACGGGAGGTTT 240
   |||||
Db 181 AATAAGCCCATTTTCCCGGCGTTTGTGATCGTAGAGGCGATGCGCAAAACGGGAGGTTT 240
   |||||

QY 241 TTAGCCCTTCACTAGCTTGTGGGGTTTGACCTCGAATCGCCAAAACAAAAATCGTGTAT 300
   |||||
Db 241 TTAGCCCTTCACTAGCTTGTGGGGTTTGACCTCGAATCGCCAAAACAAAAATCGTGTAT 300
   |||||

QY 301 TTCATGACGATGATAGGTTAAATTCGCGATCCCTGTAACCCAGCGACAGATTAGAA 360
   |||||
Db 301 TTCATGACGATGATAGGTTAAATTCGCGATCCCTGTAACCCAGCGACAGATTAGAA 360
   |||||

QY 361 TACCATTTAGAAGTCTTAAAGCATAGGGCATGATCTGGCAAGTGGGTGGCAGGCTCAA 420
   |||||
Db 361 TACCATTTAGAAGTCTTAAAGCATAGGGCATGATCTGGCAAGTGGGTGGCAGGCTCAA 420
   |||||

QY 421 GTGGATGGCAAGTGTGCTGAAGCGGAATTGAAGCCCATGATTGACAGAGAGATTAA 480
   |||||
Db 421 GTGGATGGCAAGTGTGCTGAAGCGGAATTGAAGCCCATGATTGACAGAGAGATTAA 480
   |||||

RESULT 3
US-09-815-242-7504
; Sequence 7504, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7504
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Helicobacter pylori
```

```
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(480)
US-09-815-242-7504

Query Match          94.7%; Score 454.4; DB 3; Length 480;
Best Local Similarity 96.7%; Pred. No. 1.3e-123;
Matches 464; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGGAACAAAGCCATCAAACTTGGCAATCTCAATTTTATAGAGCATATCTTACAAAT 60
DB 1 ATGGAACAAAGCCATCAAACTTGGCAATCTCAATTTTATAGAGCATATCTTACAGAT 60
QY 61 CTACCTCACCCTATCCCATCTTTTAGTGGATAGATTTATAGATTAAGCCCAATAA 120
DB 61 TTACCTCACCCTATCCCATCTTTTAGTGGATAGATTTATAGATTAAGCCCAATAA 120
QY 121 AAAATTTGCGTTTATAAGATATCACTTTTAAATGAAGACGTGTTAAACGGGCAATTCCT 180
DB 121 AAAATTTGCGTTTATAAGATATCACTTTTAAATGAAGATGTGTTAAACGGGCAATTCCT 180
QY 181 AATAAGCCCAATTTCCCGGCGTTTGTGATCGTAGAGGCATGCGCAAAACGGGAGGTTT 240
DB 181 AATAAGCCCAATTTCCCGGCGTTTGTGATCGTAGAGGCATGCGCAAAACGGGAGGTTT 240
QY 241 TTAGCCTTCACTAGCTTTGCGGGGTTTGAACCTCGTAAGTCCGCAAAACAAATAATCGTGTAT 300
DB 241 TTAGCCTTCACTAGCTTTGCGGGGTTTGAACCTCGTAAGTCCGCAAAACAAATAATCGTGTAT 300
QY 301 TTCATGACGATTTGATAGGTTAAATTCGCGATCCCTGTAAACCCAGGCGACAGATTAGAA 360
DB 301 TTCATGACGATTTGATAGGTTAAATTCGCGATCCCTGTAAACCCAGGCGACAGATTAGAA 360
QY 421 GTGGATGCAAAAGTGTGCTGCTGAAGCCGAATTAAGAGCCATGATCGCAGAGAGAGATAA 480
DB 421 GTGGATGCAAAAGTGTGCTGCTGAAGCCGAATTAAGAGCCATGATCGCAGAGAGAGATAA 480

RESULT 4
US-10-335-977-2213
; Sequence 2213, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207

;
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 2213:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...480
; SEQUENCE DESCRIPTION: SEQ ID NO: 2213:
US-10-335-977-2213

Query Match          94.7%; Score 454.4; DB 7; Length 480;
Best Local Similarity 96.7%; Pred. No. 1.3e-123;
Matches 464; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGGAACAAAGCCATCAAACTTGGCAATCTCAATTTTATAGAGCATATCTTACAAAT 60
DB 1 ATGGAACAAAGCCATCAAACTTGGCAATCTCAATTTTATAGAGCATATCTTACAGAT 60
QY 61 CTACCTCACCCTATCCCATCTTTTAGTGGATAGATTTATAGATTAAGCCCAATAA 120
DB 61 TTACCTCACCCTATCCCATCTTTTAGTGGATAGATTTATAGATTAAGCCCAATAA 120
QY 121 AAAATTTGCGTTTATAAGATATCACTTTTAAATGAAGACGTGTTAAACGGGCAATTCCT 180
DB 121 AAAATTTGCGTTTATAAGATATCACTTTTAAATGAAGATGTGTTAAACGGGCAATTCCT 180
QY 181 AATAAGCCCAATTTCCCGGCGTTTGTGATCGTAGAGGCATGCGCAAAACGGGAGGTTT 240
DB 181 AATAAGCCCAATTTCCCGGCGTTTGTGATCGTAGAGGCATGCGCAAAACGGGAGGTTT 240
QY 241 TTAGCCTTCACTAGCTTTGCGGGGTTTGAACCTCGTAAGTCCGCAAAACAAATAATCGTGTAT 300
DB 241 TTAGCCTTCACTAGCTTTGCGGGGTTTGAACCTCGTAAGTCCGCAAAACAAATAATCGTGTAT 300
QY 301 TTCATGACGATTTGATAGGTTAAATTCGCGATCCCTGTAAACCCAGGCGACAGATTAGAA 360
DB 301 TTCATGACGATTTGATAGGTTAAATTCGCGATCCCTGTAAACCCAGGCGACAGATTAGAA 360
QY 421 GTGGATGCAAAAGTGTGCTGCTGAAGCCGAATTAAGAGCCATGATCGCAGAGAGAGATAA 480
DB 421 GTGGATGCAAAAGTGTGCTGCTGAAGCCGAATTAAGAGCCATGATCGCAGAGAGAGATAA 480

RESULT 5
US-10-335-977-2212
; Sequence 2212, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
```

COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-Dec-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 2212:
SEQUENCE CHARACTERISTICS:
LENGTH: 180 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...180
SEQUENCE DESCRIPTION: SEQ ID NO: 2212:
US-10-335-977-2212

Query Match 30.3%; Score 145.4; DB 7; Length 180;
Best Local Similarity 92.2%; Pred. No. 2e-32;
Matches 165; Conservative 0; Mismatches 11; Indels 3; Gaps 1;
Qy 161 TGTTTAACGGGCAATTCCTAATAAGCCCAATTTCCCGGCGGTTTGTATCGTAGAGGGCA 220
Db 2 TGTGTATGGGCAATTCCTAATAAGCCCAATTTCCCGGCGGTTTGTATCGTAGAGGGCA 61
Qy 221 TGGCGCAACGGAGGGTTTGTAGCCTTCACTAGCTTGTGGGGGTTTGACCTCGAATCG 280
Db 62 TGGCGCAACGGAGGGTTTGTAGCCTTCACTAGCTTGTGGGGGTTTGACCTCGAATAG 121
Qy 281 CCA--AAACAAAATCGGTATTTTCATGACGATTGATAAGGTTAAATTCGCGATCCCT 336
Db 122 CCAACAAAATAATCCGTTATTTTCATGACGATTGACAGGTTAAATTCGCGATTCCT 180

RESULT 6

US-10-282-122A-18085
Sequence 18085, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carx, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITFA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 18085
LENGTH: 441
TYPE: DNA
ORGANISM: Campylobacter jejuni
US-10-282-122A-18085

Query Match 28.0%; Score 134.4; DB 7; Length 441;
Best Local Similarity 58.4%; Pred. No. 5.7e-29;
Matches 256; Conservative 0; Mismatches 176; Indels 6; Gaps 1;
Qy 39 TATAGAGCATATCTTACAAATTCCTACCTCAGCGCTATCCCATGCTTTTACTAGTAGATAAT 98
Db 9 TGTAAATGCAAAATTCAGAAATCTTACCACACCGTTATCCCTTTTACTAGTAGATAATAAT 68
Qy 99 TATAGAGTTACAGCCCAATAAAAAATTTGTCGCTTATAAGAAATATCACTTTTAATGAAGA 158
Db 69 TACAGAAATTAAGATTAAAGAGTGGTGTAGGATATATAAAACATCAGTATAAGTGACCA 128
Qy 159 CGTGTTTTAAACGGGCAATTTCCCTAATAAGCCCAATTTCCCGGCGGTTTGTATCGTAGAGG 218
Db 129 TGTTTTATGGACATTTCCAGGCGCATCTTATTTATCTCGAGTTTTTGTATTTAGAAGG 188
Qy 219 CATGGCGCAACGGGAGGGTTTTTGTAGCCTTCACTAGCTTGTGGGGGTTTGACCCCTGAAT 278
Db 189 TATGGCTCAAAACAGGTGGAGTTTTTGTAGCTTTTGAAGAT- - - - -GGAAGATAAAGTGA 242
Qy 279 CGCCAAAACAAAATCGTGTATTTTCATGACGATTGATAAGGTTAAATTCGCGATCCCTGT 338
Db 243 TCCAAAAGTAAAGTAGTTTATTTTACAGGCGATAGAGCAAAATTTAGAAATCCCTGT 302
Qy 339 AACCCAGCGCACAGATTAGAAATACCATTTAGAAGTCTTTAAAGCATAAGGGCATGATCTG 398
Db 303 CGCTCTGGAGATAGCTTGTATTAAGATGAGCGTGTGTTAAATTCGTTGTAATATGTG 362
Qy 399 GCAAGTGGGTGGCACGGCTCAAGTGGATGGCAAGTGGTTCGCTGAAGCGCAATTTGAAGC 458
Db 363 GATTTTAAAGGGCAAGCTTTTGTAGATGGAAATTTAGTTGCGAGAGCGCGAGCTTTAAAGC 422
Qy 459 CATGATTGCAGAGAGAGA 476
Db 423 TATGATAGTGGATAAATA 440

RESULT 7

US-10-369-493-35520
Sequence 35520, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 35520
; LENGTH: 459
; TYPE: DNA
; ORGANISM: Mesorhizobium loti
US-10-369-493-35520

Query Match 20.7%; Score 99.2; DB 6; Length 459;
Best Local Similarity 55.1%; Pred. No. 1.5e-18;
Matches 225; Conservative 0; Mismatches 168; Indels 15; Gaps 1;

Qy	59	TTCTACCTCACCGCTATCCCATCTTTAGTGGATAGATTATAGAGTTACAGCCCAATA	118
Db	53	TCCTGCACACCGCTATCCCTTCTGATGATCGACGCATCATCGATCGATGGTGACG	112
Qy	119	AAAAAATTGCTGCTTATAAGAAATATCACTTTTAAATGAAGACGTGTTTAAACGGGCATTTCC	178
Db	113	ATTCGCCCATCGGCATCAAGACGTCAACATCAAGCGCGCATTTTCAGGGCATTTCC	172
Qy	179	CTAATAAGCCCAATTTTCCGGGCGTTTGTATCGTAGAGGCATGGCGCAACCGGAGGGT	238
Db	173	CGGAGCAGCGGCGTATGCGCGGGCGTGTGATCGTGAAGCCATGCGCAGACGGCAGCG	232
Qy	239	TTTTAGCCTTCACTAGCTTGTGGGGGTTTGAACCTGMAATCGCAAAACAAAATCGTGT	298
Db	233	CCATCTGCATCGCAGCCTCGGGCGT-----CGAAGCGCTCGCTGCTCT	277
Qy	299	ATTTCATGACGATTGATAAGTTTAAATTCGCATCCTGTAACCCAGCGCAGAGATTAG	358
Db	278	ATTTCGACCATCGACACGCGCAATTCGCAACCGTGTTCGCGGCGACAGTTGA	337
Qy	359	AATACCAATTAGAGTCTTAAAGCATAAGCGGCATGATCTGGCAAGTGGGTGGCAGCGCTC	418
Db	338	AGATTCACTCAAAAAATCAAGACGCGGCAACCTGCTCAAAATTCGCTGTGAAGCCC	397
Qy	419	AAGTGATGGCAAAAGTGGTGGCTGGAAGCGCAATTTGAAGCCATGATTG	466
Db	398	TGGTCGATGGCACCAGGCGGCGGCGAGATTTTCAGCCATGATGG	445

RESULT 8

US-10-282-122A-16557
; Sequence 16557, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20

RESULT 9

US-10-282-122A-21741
; Sequence 21741, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos

; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16557
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Clostridium botulinum
US-10-282-122A-16557

Query Match 19.8%; Score 95; DB 7; Length 432;
Best Local Similarity 53.5%; Pred. No. 2.5e-17;
Matches 234; Conservative 0; Mismatches 185; Indels 18; Gaps 1;

Qy	34	TTTTTTATAGAGCATATCTTACAAATTTCTACTCACCGCTATCCCATGCTTTTAGTGAT	93
Db	13	TTAGATATAATGAATAAAAAAATTTATCCGCACAGATATCCATTTTATTAGTAGT	72
Qy	94	AGAAATATAGAGTTTACAAGCCAATAAAAAAATTTGCTGCTTATAAGAAATATCATTTTAAAT	153
Db	73	AAAATACTCAATTAGAAGAGGGAAGTGCAGTTGGATATAAAAAATGTTACAGCTAAT	132
Qy	154	GAAGACGTGTTTAAACGGGCATTTCCCTAAATAAGCCCATTTTCCGGGGGCTTTTGATCGTA	213
Db	133	GAATACCTTTTAAATGGGCATTTTCCGGAAGAACAGTAATGCTCGGGGTTCTGATTATA	192
Qy	214	GAGGCATGCGCAACGCGGAGGGTTTGTAGCCCTTCACTAGCTTGTGGGGGTTTGACCCCT	273
Db	193	GAAGCTTTTAGCACAGGTGCGAGCTGTTGCTATTTTAAAGTA-----AA	234
Qy	274	GAATCGCCAAACAAAAATCGTGTATTTTCATGACGATTGATAGGTTAAATTCGCGCATC	333
Db	235	GAAGAGTTTAAAGGAATAATAGCTTATTTTGGAGGTATAAATAAGCTAAGTTTCAAAAA	294
Qy	334	CTGTAAACCCAGCGCAGATAGATAATACATTTAGAAGTCTTAAAGCATAAAGGGCATG	393
Db	295	AAGGTAGTACACAGAGATGTTTAAAACTTAGTAGTAGATCTTACTTAAATAAAGGCGTT	354
Qy	394	ATCTGGCAAGTGGTGGCAGCGCTCAAGTGGTGAAGTGGTGGTGAAGCCGAATTG	453
Db	355	GCAGGAGTAGGTAAGGCCGTAGCTACTGTAGATCGGAAAGTAGTGTGCAGNAGCAGATTA	414
Qy	454	AAAGCCATGATTGCAGA	470
Db	415	TTATTTGTAATAGGAAA	431

	Query Match	19.5%;	Score 93.4;	DB 7;	Length 417;
	Best Local Similarity	53.6%;	Pred. No. 7.4e-17;		
	Matches 230;	Conservative 0;	Mismatches 181;	Indels 18;	Gaps 1;
Qy	40	ATAGAGCATATCTTACAAATCTACCTCACCGCTATCCATGCTTTTAGTGGATAGAATT	99		
Db	7	ATACAAGAAATAAAGAAATCAATCCACATCGTTATCCAAATGCTTTTGGATTGATCGAGTG	66		
Qy	100	ATAGAGTTCAAGCCCAATAAAAAATTTGTCCTTATAAGAAATACACTTTTAATGAAGAC	159		
Db	67	GAAGAAATGTTGAAGAGAAACGAGTGTGCTAAGAAATAATGAACATCATGAGCCC	126		
Qy	160	GTGTTTAAACGGGCATTTCCCTAAATAAGCCCAATTTCCCGGCGTTTTGATCGTAGAGGGC	219		
Db	127	TTTTTTCAAGGACATTTTCCAGAAGAACCTGTTATGCCAGGGTATTGATTGTGAAGCA	186		
Qy	220	ATGGCGGAAACGGGAGGGTTTTTAGCCTTCACTAGCTTGTGGGGGTTTGACCTGAAATC	279		
Db	187	ATGGCTCAAGCAGGAGCAGTTGCATTACTTTCT-----CTTGAACAA	228		
Qy	280	GCCAAACAAAAATCGTGTAATTTTCATGACGATTGATAAGTTTAAATTCGCATCCCTGTGA	339		
Db	229	TTCAAGGGAAAGACGGGTATTATTCGGCGGATTTGACAAAGCAAAATTTCCGTAGAAGAGTT	288		
Qy	340	ACCCGAGCGACAGATTAGAAATACATTTAGAAGCTTTAAAGCATAAAGGCGCATGATCTGG	399		
Db	289	ACACGGGAGATACCTTATACCTTTGAAGTTGAGATATTGAAGTGAAGCTTCAGCCGGA	348		
Qy	400	CAAGTGGTGGCACGGCTCAAGTGGATGGCAAGATGGTCGTGAAGCGCGAATTGAAAGCC	459		

	Query Match	19.3%;	Score 92.6;	DB 7;	Length 462;
	Best Local Similarity	53.8%;	Pred. No. 1.4e-16;		
	Matches 222;	Conservative	0;	Mismatches 179;	Indels 12; Gaps 1;
Qy	59	TTCTACCTCACCGCTATCCATGCTTTTACGTGGTAGAATATATAGAGTTACACGCCAATA	118		
Db	56	TTCTGCGCTACCGGTATCCATTTTACTGANTGATCGCGTGACCGACTATGAAGAAGGTA	115		
Qy	119	AAAAAATGTGCGCTTATAAGAAATACACTTTTAAATGAAGACGTCTTTTAAACGGGCAATTC	178		
Db	116	AGTACCTGATGTGGTTTCAAAAAGTTTTCAGTGANTGAGCCCTCAGTTTACGGGTACATTTC	175		
Qy	179	CTAATAAGCCCATTTTCCCGGGCGTTTGTATCGTAGAGGGCATGGCGCAACCGGAGGGT	238		
Db	176	CACAACTGCCGGTTTTCCCGGTGTAATGATTTTGAAGGCGATGGCTCAAGCCACTGGTC	235		

Qy	239	TTTTAGCCTTCAC	TAGCTTGTGGGGT	TGACCCCTGAAATCGCCAAA	CAAAAATCGTGT	298
Db	236	TGTTGGCGTTCAAA	ACCTTGTGTGCGCCAA	-----	AGAGATGAATCTGACT	283
Qy	299	ATTTTCATGACGAT	TGATAAGGTTAAAT	TTCGCGATCCCTGTAACCC	CAGCGCACAGATTAG	358
Db	284	ATTTTGCACGAT	TGATGAGCTAAGTT	TCGTAAGCCAGTGA	CACGGGCGAGTCAACTGA	343
Qy	359	AATAACCATTTAGA	AGTCTTAAAGCA	TAAAGGGCATGATCTTGGCAAG	TGGGTGGCACGGCTC	418
Db	344	TGGTTGAATGTAAT	TCTTAAAGAGCGT	CGTGGTATCGCGTGT	TTAACGGCGTGCCA	403
Qy	419	AAGTGGATGGCAAA	GTGTGCTCGCTGAG	CGCGAATTGAAAGCC	ATGATGTCAGAG	471
Db	404	AAGTCGATGTGTG	TGTGTGTCTGCTCAACT	GAAATGCGCACG	CTCGAGAG	456

```

RESULT 11
US-10-282-122A-24354
; Sequence 24354, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

```

Query Match 18.5%; Score 88.6; DB 7; Length 435;
Best Local Similarity 64.3%; Pred. No. 2e-15;
Matches 133; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 34 TTTTATTAGAGCATATCTTACAAATCTCTACCTACCGCTATCCCATGCTTTTAGTGAT 93

```

Db      4   TTAGATATTAGAAAAATCAAGAGAGATTTTGCTCATCTGTTATCAATTTTGTGTAGTAGAT   63
Qy      94   AGAATTATAGAGTTTACAGGCCAATAAAAAAATGTGCTTATTAAGAAATACACTTTTAAAT   153
Db     164   AGAGTTATTTCTGTTGAAGAGGCCAAAAAGTTACCGCTATTAGAAATGTAACAGCCAAAT   123
Qy     154   GAAGACGTGTTTAAACGGGCATTTCCCTTAATAAAGGCCAATTTTCCCGGGCGTTTTGTGATCGTA   213
Db     124   GAAGAAATTTTTTAAACGGGCACATTTTCCCTCAGTATCTCTGTAATGCCAGGAGTATTAATAGTG   183
Qy     214   GAGGCGCATGGCGCAACAGCGGAGGGTTT   240
Db     184   GAAGCTCTGGCGCAACAAAGTGGAAAT   210

RESULT 12
US-10-398-221-1922
; Sequence 1922, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; CURRENT APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1922
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Listeria monocytogenes-4B
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide a,g,c or t/u
US-10-398-221-1922

```

; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3523
; LENGTH: 1927
; TYPE: DNA
; ORGANISM: Listeria monocytogenes 4b
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u

US-10-398-221-3523

Query Match 18.5%; Score 88.6; DB 7; Length 1927;
Best Local Similarity 64.3%; Pred. No. 4.4e-15;
Matches 133; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 34 TTTTATAGACATCTTACAAATTTCTACCTCACCGCTATCCCATGCTTTTAGTGGAT 93
Db 990 TTAGATATTAGAAATCAAGAGATTTTGCTCATCGTTATCCATTTTGTAGTAGAT 1049
QY 94 AGAATTTATAGAGTTACAAGCCCAATAAAAAATTTGCTTTATAAGAATATCACTTTTAAT 153
Db 1050 AGAGTTATTCTGTTGGAAGAGCAAAAAATTTACCGCTATTAAAGATGTACAGCAAT 1109
QY 154 GAAGAGTGTTTAAAGCGGCAATTTCCCTAATAAGCCCAATTTCCCGGCGCTTTTATCGTA 213
Db 1110 GAAGAATTTTAAAGCGGCACTTTCTGAGTATCTGTATATGCCAGGAGTTTAAATAGTG 1169
QY 214 GAGGCGATGCGCGCAACGGAGGGTTT 240
Db 1170 GAAGCTCTAGCGCAAACTAGTGGAAAT 1196

RESULT 14

US-10-282-122A-31282
; Sequence 31282, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31282
; LENGTH: 459
; TYPE: DNA
; ORGANISM: Pasteurella multocida
; US-10-282-122A-31282

Query Match 18.4%; Score 88.2; DB 7; Length 459;
Best Local Similarity 53.3%; Pred. No. 2.7e-15;
Matches 220; Conservative 0; Mismatches 178; Indels 15; Gaps 1;
QY 42 AGAGCATATCTTACAAATTTCTACCTCACCGCTATCCCATGCTTTTAGTGGATAGATAT 101
Db 39 AAATGAAATTTATGAAATTTGCTCTCACCGTTATCCTTTTATTTAGTGGATCGTGTGGT 98
QY 102 AGATTACAGCAATAAAAAATTTGCTTATAGCAATATCACTTTTAATGAAGCGT 161
Db 99 GGATTATGAAGAGGTAAATGGCTTAAAGCTGTAAAAAATATTAGCGTAAATGAGCCTTG 158
QY 162 GTTTAACGGGCATTTCCCTAATAAGCCCAATTTTCCCGGGCTTTTGTATCGTAGAGGCAT 221
Db 159 CTTTACGGGACATTTCCGAGCAACACGATTTTCCGGGTGTGTGATTTTGGAGCCAT 218
QY 222 GCGCAACCGGGAGGGTTTTAGCCTTCTACCTAGCTTGTGGGGTTTGACCTGAAATCGC 281
Db 219 GGCACAAGCGAGCGGGGTGCTCGCCTGTAAAAAATTACCGTCAATTAG----- 265
QY 282 CAAAAAATAATCGTGTATTTTCATGACGATGATAGGTTAAATTCGCACTCCCTGTAAAC 341
Db 266 --AAATGAAATTTATTTATTTTGTGCAATTTGATATGCGCGCTTTAAACGCTCTGTGTT 323
QY 342 CCCAGGCGACAGATTAGAAATACCAATTTAGAGTCTTTAAAGCATAAGGCGCATGATCTGCA 401
Db 324 GCCTGGTGATCAAAATGGTGTAGAGTTCACTTCTTGAAAGAGCGTCCGGGTATTACGG 383
QY 402 AGTGGGTGGCACCGGCTCAAGTGATGGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 454
Db 384 TTTCACTGGTGTGCGCAACCGTAGATGCTCAAGTGGTGTGCGAAGCAGAAATTAA 436

RESULT 15

US-10-282-122A-14857
; Sequence 14857, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14857
; TYPE: DNA
; LENGTH: 453
; ORGANISM: Bordetella pertussis
US-10-282-122A-14857

Query Match 18.2%; Score 87.4; DB 7; Length 453;
Best Local Similarity 50.4%; Pred. No. 4.6e-15;
Matches 214; Conservative 0; Mismatches 211; Indels 0; Gaps 0;

QY	40	ATAGAGCATATCTTACAAATTTACCTACCGCTATCCCATGCTTTTAGTGGATAGAAAT	99
DB	13	ATCAAGGGATCATGGACCGGCTGCCGCATCGCTACCCGATGCTGATCGACCGGTG	72
QY	100	ATAGAGTTACAAGCCAATAAATAATTTGCGCTTATAAGAAATATCACTTTTAATGAAGAC	159
DB	73	CTGGAGATGGTCCCGGGAATCCATCGTTGCCATCAAGAACGTCTCGATCAACGAGCCG	132
QY	160	GTGTTTAAACGGGCAATTCCTTAATAAGCCCATTTTCCCGGGCGTTTGTATCGTAGAGGC	219
DB	133	TTTTTCACGGGCCATTTTCGGCATCATCGGTATCGCGGGCGTGTGATCGTCGAGGCC	192
QY	220	ATGGCGCAACCGGAGGGTTTTTAGCTTTCACTAGCTTGTGGGGGTTTGACCCCTGAAATC	279
DB	193	ATGGCGCAGGCTTCGGCCCTGTTCTCGTTTCAACGACGAGAAACGGCGGGCTGAAGTCGCAC	252
QY	280	GCCAAACAAAATCGTGTATTTTCATGACGATTGATAAGTTAAATTCGCATCCCTGTA	339
DB	253	GGCGCCAAACCGCCCTACTACCTGTGGGCATCGACGGCGCGCGTTCCGTAAACCCGTG	312
QY	340	ACCCAGGCGACAGATTAGAATACCATTTAGAAAGTCTTAAAGCATAAAGGCGATGATCTGG	399
DB	313	GTGCGGGGACACAGTTTGGCGGTGGNAGTCGAGGCCGAGGCCGTGCGGTACCATCTGC	372
QY	400	CAAGTGGGTGGACCGGCTCAAGTGGATGGCAAGTGGTCGCTGAAGCCGGAATTTGAAAGCC	459
DB	373	AAGTACCGGGCGGGCCCTGTTGTCGACGGCCAACTGGTCCGCGAAGCCAAAGCTGATGTC	432
QY	460	ATGAT 464	
DB	433	GCAT 437	

Search completed: January 12, 2006, 03:25:50
Job time : 659 secs

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New). Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 01:29:32 ; Search time 226 Seconds
(without alignments)
1718.970 Million cell updates/sec

Title: US-10-662-126-36

Perfect score: 480

Sequence: 1 atggaacaaagccatcaaaa.....tgattgcagagagattaa 480

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 6038814 seqs, 404674181 residues

Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.New.*

- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
- 10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73.6	15.3	447	6	US-10-467-657-6109 Sequence 6109, Ap
2	73.6	15.3	447	6	US-10-467-657-7549 Sequence 7549, Ap
3	73.6	15.3	552	6	US-10-467-657-6129 Sequence 6129, Ap
c 4	73.6	15.3	552	6	US-10-467-657-7559 Sequence 7559, Ap
c 5	60.8	12.7	2926	6	US-10-793-626-3747 Sequence 3747, Ap
6	35.8	7.5	1458	6	US-10-750-185-26969 Sequence 26969, A
7	35.8	7.5	1458	6	US-10-750-623-26969 Sequence 26969, A
8	35.6	7.4	822	6	US-10-793-626-727 Sequence 727, App
9	35.6	7.4	3215	6	US-10-793-626-3566 Sequence 3566, Ap
c 10	35.6	7.4	3993	6	US-10-793-626-3985 Sequence 3985, A
c 11	34.2	7.1	2528	6	US-10-750-185-51435 Sequence 51435, A
c 12	34.2	7.1	2528	6	US-10-750-623-51435 Sequence 51435, A
c 13	34	7.1	134499	7	US-11-117-187-192 Sequence 192, App
14	33.6	7.0	2382	6	US-10-750-185-57788 Sequence 57788, A
15	33.6	7.0	2382	6	US-10-750-623-57788 Sequence 57788, A
c 16	33.6	7.0	394468	7	US-10-995-561-13473 Sequence 13473, A
17	33.4	7.0	856	6	US-10-750-185-60462 Sequence 60462, A
18	33.4	7.0	856	6	US-10-750-623-60462 Sequence 60462, A
c 19	33.4	7.0	1304	6	US-10-750-185-52022 Sequence 52022, A
c 20	33.4	7.0	1304	6	US-10-750-623-52022 Sequence 52022, A
c 21	33.4	7.0	6182	6	US-10-240-708-87 Sequence 87, Appl
c 22	33	6.9	555	7	US-11-128-061-7215 Sequence 7215, Ap
c 23	33	6.9	2372	7	US-11-128-061-3573 Sequence 3573, Ap

ALIGNMENTS

RESULT 1

US-10-467-657-6109 7 600 6.8 32.8 6.8 600 7 US-11-136-527-5665 Sequence 5665, Ap
; Sequence 6109, Application US/10467657 Sequence 39282, A
; Publication No. US20050260581A1 Sequence 39282, A
; GENERAL INFORMATION: Sequence 1569, Ap
; APPLICANT: CHIRON Spa Sequence 53522, A
; APPLICANT: FONTANA Maria Rita Sequence 53522, A
; APPLICANT: PIZZA Mariagrazia Sequence 396, App
; APPLICANT: MASIGNANI Vega Sequence 59973, A
; APPLICANT: MONACI Elisabetta Sequence 59973, A
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS Sequence 51969, A
; FILE REFERENCE: Sequence 51969, A
; CURRENT APPLICATION NUMBER: US/10/467,657 Sequence 55313, A
; CURRENT FILING DATE: 2003-08-11 Sequence 13264, A
; PRIOR APPLICATION NUMBER: GB-0103424.8 Sequence 31, Appl
; PRIOR FILING DATE: 2001-02-12 Sequence 3663, A
; NUMBER OF SEQ ID NOS: 9218 Sequence 47750, A
; SOFTWARE: SeqWin99, version 1.04 Sequence 47750, A
; SEQ ID NO 6109 Sequence 3384, Ap
; LENGTH: 447 Sequence 32382, A
; TYPE: DNA Sequence 13293, A
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6109

Query Match	15.3%	Score 73.6	DB 6	Length 447
Best Local Similarity	51.2%	Pred. No. 1.7e-13		
Matches	206	Conservative	0	Mismatches 184
				Indels 12
				Gaps 1
Qy	47	ATATCTTACAAATCTACCTCACCGCTATCCATGCTTTTAGTGGATAGAATTATAGAGT	106	
Db	32	ACATCCAAAACCTATCCCGCTACCGCTTCTCCAGCTCGACCGCATACCGCT	91	
Qy	107	TACAAGCCAATAAAAAAATGTCTGCTTATAGAATATACATTTTAAAGAGCGTGTFTA	166	
Db	92	TCGAGCGCATGMAAACCTGACCGCCATCAAAAAGTAACCAACGACCCCAATTC	151	
Qy	167	ACGGGCAATTTCCCTAATAAGCCCATTTTCCGGGGGTTTTCATCTAGAGGCGATGGCG	226	
Db	152	AAGGCCATTTCCCGGACCTGCGGTTATGCGCGGTACTCATCATCGAAGCGATGGCG	211	
Qy	227	AAACGGGAGGTTTATAGCCTTCACTAGCTTGTGGGGTTTGACCTGAAATCGCAAAA	286	
Db	212	AGCGTGGCGACGTTTGGCGATTTTGACGGAAGCGGGG	259	
Qy	287	CAAAAATCGTGTATTTTCATGACGATTGATAGGTTAAATTCGCGATCCCTGTAAACCCAG	346	

Db 260 ACGAATTTTCTTCTTCGCGCGCATAGACGAAGCCGTTTCAAACGCCAAGTCATCCCCG 319
Qy 347 GCGACAGATTAGAAATACCAATTTAGAAAGTCTTAAAGCATAAAGGCGCATGATCTGGCAAGTGG 406
Db 320 GCGACCAACTCTCTTTGAAGTCGAATCTCTGACNAGCGCGCGGCATCGGCAAAATTC 379
Qy 407 GTGGCAGCGCTCAAGTGGATGGCAAAAGTGGTGCCTGAAGCCG 448
Db 380 ACGCCGTTGCCAAAGTGGACGACGAAGTGCCTGCGAAGCCG 421

RESULT 2

US-10-467-657-7549
; Sequence 7549, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7549
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7549

Query Match 15.3%; Score 73.6; DB 6; Length 447;
Best Local Similarity 51.2%; Pred. No. 1.7e-13;
Matches 206; Conservative 0; Mismatches 184; Indels 12; Gaps 1;
Qy 47 ATATCTTACAAATCTACCTACCGCTATCCCATGCTTTTAGTGATAGAAATATAGAGT 106
Db 32 ACATCAAAAACCTCATCCGCCACCGCTACCGCTTCTCCAGCTCGACCGCATTACCGCCT 91
Qy 107 TACAAGCCCAATAAAAAATTTGCGCTTTATAAGAAATATCACTTTTAATGAAGAGTGTTTA 166
Db 92 TCGAGCCGATGAACCCCTGACCGCATCAAAAACGTAACCAATAAGAACCCCAATTC 151
Qy 167 ACGGGCAATTTCCCTAATAAGCCCAATTTTCCCGGGCGGTTTGTAGTGGGCGCATGCGCG 226
Db 152 AAGGCCATTTCCCGACCTGCGCGTTATGCGCGCGTACTCATCATCGAAGCGATGCGCG 211
Qy 227 AAACGGAGGGTTTGTAGCTTCACTAGCTTGTGGGGGTTTGAACCTCGTAAGAAATCGCCAAA 286
Db 212 AGCGGTGCGGCACGTTGGCGATTTTGAGCGAAGCGCGG-----CGCAAGGAAA 259
Qy 287 CAAAATCGTGATTTTCATGACCAATGATAGTAAATTCGCCATCCCTGTAAACCCAG 346
Db 260 ACGAATTTTCTTCTTCGCGGCATAGACGAACCGCGTTTCAAACGCCCAAGTCATCCCCG 319
Qy 347 GCGACAGATTAGAAATACCAATTTAGAAAGTCTTAAAGCATAAAGGCGCATGATCTGGCAAGTGG 406
Db 320 GCGACCAACTCGTCTTTGAAGTCGAATCTCTGACCGCGCGCGGCATCGGCAAAATTC 379
Qy 407 GTGGCAGCGCTCAAGTGGATGGCAAAAGTGGTGCCTGAAGCCG 448
Db 380 ACGCCGTTGCCAAAGTGGACGACGAAGTGCCTGCGAAGCCG 421

RESULT 3

US-10-467-657-6129
; Sequence 6129, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:

; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6129
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6129

Query Match 15.3%; Score 73.6; DB 6; Length 552;
Best Local Similarity 51.2%; Pred. No. 2e-13;
Matches 206; Conservative 0; Mismatches 184; Indels 12; Gaps 1;
Qy 47 ATATCTTACAAATCTACCTACCGCTATCCCATGCTTTTAGTGATAGAAATATAGAGT 106
Db 47 ACATCAAAAACCTCATCCGCCACCGCTACCGCTTCTCCAGCTCGACCGCATTACCGCCT 106
Qy 107 TACAAGCCCAATAAAAAATTTGCGCTTTATAAGAAATATCACTTTTAATGAAGAGTGTTTA 166
Db 107 TCGAGCCGATGAACCCCTGACCGCATCAAAAACGTAACCAATAAGAACCCCAATTC 166
Qy 167 ACGGGCAATTTCCCTAATAAGCCCAATTTTCCCGGGCGGTTTGTAGTGGGCGCATGCGCG 226
Db 167 AAGGCCATTTCCCGACCTGCGCGTTATGCGCGGTAATCATCATCGAAGCGATGCGCG 226
Qy 227 AAACGGAGGGTTTGTAGCTTCACTAGCTTGTGGGGGTTTGAACCTCGTAAGAAATCGCCAAA 286
Db 227 AGCGGTGCGGCACGTTGGCGATTTTGAGCGAAGCGCGG-----CGCAAGGAAA 274
Qy 287 CAAAATCGTGATTTTCATGACGATTCATAGAGTTAAATTCGCCATCCCTGTAAACCCAG 346
Db 275 ACGAATTTTCTTCTTCGCGGCATAGACGAACCGCGTTTCAAACGCCCAAGTCATCCCCG 334
Qy 347 GCGACAGATTAGAAATACCAATTTAGAAAGTCTTAAAGCATAAAGGCGCATGATCTGGCAAGTGG 406
Db 335 GCGACCAACTCGTCTTTGAAGTCGAATCTCTGACCGCGCGCGGCATCGGCAAAATTC 394
Qy 407 GTGGCAGCGCTCAAGTGGATGGCAAAAGTGGTGCCTGAAGCCG 448
Db 395 ACGCCGTTGCCAAAGTGGACGACGAAGTGCCTGCGCGTCGAAGCCG 436

RESULT 4

US-10-467-657-7559/c
; Sequence 7559, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7559
; LENGTH: 552
; TYPE: DNA


```
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7559

Query Match      15.3%; Score 73.6; DB 6; Length 552;
Best Local Similarity 51.2%; Pred. No. 2e-13;
Matches 206; Conservative 0; Mismatches 184; Indels 12; Gaps 1;

QY 47 ATATCTTACAAATCTTACCTCAGCGCTATCCCTGCTTTTAGTGGATAGATATAGT 106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 506 ACATCAAAATCTATCCCCACCGCTACCGCTTTCTCCAGCTCGACCGCATACCGCCT 447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 107 TACAAGCCATAAAAAATTTGTCCTTATAGATATCACTTTTAATGAAGAGTGTTTA 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 446 TCGAGCGGATGAACCTGACCGCATCAAAAAGTAACCATATACGAAACCCCAATCC 387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 167 ACGGCAATTCCTATAAAGCCATTTTCCCGGCGTTTGTAGTGTAGAGGCGATCGCGC 226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 AAGGCCATTTCCCGCGCTCCCGTTATGCGCGGCTACTCATCTGAAGCGATGCGCGC 327
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 227 AAACGGGAGGTTTTAGCTTCACTAGCTGTGGGGTTTGACCTGAAATCGCCAAA 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 AGCGGTGCGGCACGTTGGCGAATTTTGAGCGAAGCGCGG-----CGCAAGGAAA 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 287 CAAATAATCGTGATTTTCATGACGATTGATAAGGTAAATTTCCGCATCCCTGTAAACCCAG 346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 278 ACGAATTTTCTTCTTCCGCGCATAGACGAGCCGTTTCAACGCGCAAGTATCCCG 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 347 GCGACAGATTAGAAATACCAATTTAGAAAGTCTTAAAGCATAGGCGCATGATCGCAAGTGG 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 218 GCGACCAACTCGTCTTGAAGTGAAGTCTCTGACCGCGCGGCGCATCGGCAAAATTC 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 407 GTGGCAGCGCTCAAGTGATGGCAAGTGTGCTGCTGAAGCGG 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 ACGCCGTTGCAAGTGGACGAGCAAGTGTGCGCGTGAAGCGG 117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
US-10-793-626-3747/c
; Sequence 3747, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3747
; LENGTH: 2926
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3747

Query Match      12.7%; Score 60.8; DB 6; Length 2926;
Best Local Similarity 56.5%; Pred. No. 6.1e-09;
Matches 113; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 18 AAACCTGCAATCTCAATTTTATAGCATATCTTACAAATCTACCTCAGCGCTATCC 77
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 AAAATGGAACGATTTTGTATATATCAATTAACAAATATCCCATAGACAAC 141
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 78 CATGCTTTTAGTGGATAGATATAGATTACAAGCCAATAAAAAATTTGCTGCTTATA 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 140 TTTTCTACTAAGATAAAATTTGATGATGATGAGAGGTAAGAGTGTGTAGGACTAA 81
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 138 GAATATCACTTTTAATGAAGAGCGTTTAAACGGGCAATTTCCCTTAATGAACCCATTTCC 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
; ORGANISM: Bovine 19866881740741
US-10-750-623-26969

Query Match          7.5%; Score 35.8; DB 6; Length 1458;
Best Local Similarity 55.1%; Pred. No. 0.36;
Matches 70; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 25 CAATCTCAATTTTTATAGAGCATATCTTACAAATCTACCTACCGGTATCCCATGCTT 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 572 CTAAGTCCCTGTTTTAAATACGTGTGCAAAAGATTTTCCATATCTCTTGATTGGCTT 631
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 85 TTAGTGGATAGATTATAGATTACAGCCATAAAAAATTTGCTTTATAGATATC 144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 632 CTAATGAATAAAATTTTGTGAGTGAAGCCCTATAAAATAAATGTTTAATAATAATTTGCC 691
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 145 ACTTTTA 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 692 ATTTTAA 698
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-10-793-626-727
; Sequence 727, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 727
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-727

Query Match          7.4%; Score 35.6; DB 6; Length 822;
Best Local Similarity 50.0%; Pred. No. 0.3;
Matches 89; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 8 AAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATCTACCTC 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 AAAGTGTGTAATAATAAACTGCATAATTTAAGACAGAAAAATATTGAAATTTGACTG 508
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 68 ACCGCTATCCCATGCTTTTAGTGGATAGATTATAGATTACAGCCATAAAAAATTTG 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 AAGATGTTGCAACGAGTTTCCAAAATAGTGTGTAGAAAGTTTAAACCTATAAAGCTATTC 568
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 128 TCGCTTATAAGATATACATCTTTTAATGAAGACGCTGTTTAAACGGGCATTTCCCTAATAA 185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 569 ATGCTTGTAAACTTATAATGTTAATCGCTTAATGTTGCAGGTGGTGTGCTAGTAA 626
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
US-10-793-626-3566
; Sequence 3566, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3566
; LENGTH: 3215
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3566

Query Match          7.4%; Score 35.6; DB 6; Length 3215;
Best Local Similarity 50.0%; Pred. No. 0.66;
Matches 89; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 8 AAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATCTACCTC 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110 AAAGTGTGTAATAATAAACTGCATAATTTAAGACAGAAAAATATTGAAATTTGACTG 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 68 ACCGCTATCCCATGCTTTTAGTGGATAGATTATAGATTACAGCCATAAAAAATTTG 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 170 AAGATGTTGCAACGAGTTTCCAAAATAGTGTGTAGAAAGTTTAAACCTATAAAGCTATTC 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 128 TCGCTTATAAGATATACATCTTTTAATGAAGACGCTGTTTAAACGGGCATTTCCCTAATAA 185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 230 ATGCTTGTAAACTTATAATGTTAATCGCTTAATGTTGCAGGTGGTGTGCTAGTAA 287
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-10-793-626-3985/c
; Sequence 3985, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3985
; LENGTH: 3993
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3985

Query Match          7.4%; Score 35.6; DB 6; Length 3993;
Best Local Similarity 50.0%; Pred. No. 0.75;
Matches 89; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 8 AAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATCTACCTC 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1449 AAAGTGTGTAATAATAAACTGCATAATTTAAGACAGAAAAATATTGAAATTTGACTG 1390
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 68 ACCGCTATCCCATGCTTTTAGTGGATAGATTATAGATTACAGCCATAAAAAATTTG 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1389 AAGATGTTGCAACGAGTTTCCAAAATAGTGTGTAGAAAGTTTAAACCTATAAAGCTATTC 1330
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 128 TCGCTTATAAGATATACATCTTTTAATGAAGACGCTGTTTAAACGGGCATTTCCCTAATAA 185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1329 ATGCTTGTAAACTTATAATGTTAATCGCTTAATGTTGCAGGTGGTGTGCTAGTAA 1272
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-10-750-185-51435/c
; Sequence 51435, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
```

APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 51435
LENGTH: 2528
TYPE: DNA
ORGANISM: Bovine 19866880723565
US-10-750-185-51435

Query Match 7.1%; Score 34.2; DB 6; Length 2528;
Best Local Similarity 50.3%; Pred. No. 1.6;
Matches 84; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 18 AAACCTGCAATCTCAATTTTATAGAGCATATCTTACAAATCTACCTCACCGCTATCC 77
DB 323 ACACAGGAACCTCACCATTTCACATGCATAGGCTATCAACAAAACCTGCCATTGAAC 264
QY 78 CATGCTTTAGTGGATAGAAATTATAGATTACAGGCAATATAAAAAATTTGTCGCTTATAA 137
DB 263 ATAGCTGTGTTGAGTCAAGGAAGCCAGTCAAGCAACTATAATTAATATAGTTGAACA 204
QY 138 GAATATCACTTTTAATGAAGACGTGTTTAACGGGCAATTTCCCTTAATA 184
DB 203 CAAACACATCTTTAATAATACATTTAGCATCAAACTACCCCTAATA 157

RESULT 12
US-10-750-623-51435/c
Sequence 51435, Application US/10750623
Publication No. US20050287531A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-1
CURRENT APPLICATION NUMBER: US/10/750,623
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 51435
LENGTH: 2528
TYPE: DNA
ORGANISM: Bovine 19866880723565
US-10-750-623-51435

Query Match 7.1%; Score 34.2; DB 6; Length 2528;
Best Local Similarity 50.3%; Pred. No. 1.6;
Matches 84; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 18 AAACCTGCAATCTCAATTTTATAGAGCATATCTTACAAATCTACCTCACCGCTATCC 77
DB 323 ACACAGGAACCTCACCATTTCACATGCATAGGCTATCAACAAAACCTGCCATTGAAC 264
QY 78 CATGCTTTAGTGGATAGAAATTATAGATTACAGGCAATATAAAAAATTTGTCGCTTATAA 137

Db 263 ATAGCTGTGTTTCAAGGAAAGCCAGTCAGGACAACTATTAAATTTATAGTTGAACA 204
QY 138 GAATATCACTTTTAAATGAAGACGTGTTTAACGGGCAATTTCCCTTAATA 184
DB 203 CAAACACATCTTTAATAATACATTTAGCATCAAACTACCCCTAATA 157

RESULT 13
US-11-117-187-192/c
Sequence 192, Application US/11117187
Publication No. US20050266560A1
GENERAL INFORMATION:
APPLICANT: PREUSS, DAPHNE
APPLICANT: COPENHAVER, GREGORY
TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
FILE REFERENCE: ARCD:309US
CURRENT APPLICATION NUMBER: US/11/117,187
CURRENT FILING DATE: 2005-04-28
PRIOR APPLICATION NUMBER: US/09/531,120
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/125,219
PRIOR FILING DATE: 1999-03-18
NUMBER OF SEQ ID NOS: 212
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 192
LENGTH: 134499
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-11-117-187-192

Query Match 7.1%; Score 34; DB 7; Length 134499;
Best Local Similarity 50.6%; Pred. No. 19;
Matches 82; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 22 TTGCAATCTCAATTTTATAGAGCATATCTTACAAATTTCTACCTCACCGCTATCCCATG 81
DB 53102 TTTCTAGGTTCATTTGTATGTCTCATAGATAACAAATCTCTAATCTCTGCTTAATCAAG 53043
QY 82 CTTTTGTAGTGATAGAAATTATAGAGTTACAGCCCAATATAAAAAATTTGTCGCTTATAAGAA 141
DB 53042 TTAAATTTGGCGTTAAATATCAAAATTTATAGAAATATAAAACAATTTACAGTTATATCT 52983
QY 142 ATCACTTTTAAATGAAGACGTGTTTAACGGGCAATTTCCCTAAT 183
DB 52982 ATTACATATAATTTAGATTCACTCAACATGATGTCACCTTAT 52941

RESULT 14
US-10-750-185-57788
Sequence 57788, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 57788
LENGTH: 2382
TYPE: DNA
ORGANISM: Bovine 19866881295850
US-10-750-185-57788

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 03:12:12 ; Search time 3014 Seconds
(without alignments)
9052.706 Million cell updates/sec

Title: US-10-662-126-36
Perfect score: 480
Sequence: 1 atggaacaaagccatcaaaaa.....tgattcgagagagattaa 480

Scoring table: OLIGO NUC
Gapop 60.0 , Capext 60.0
Searched: 5883141 seqs, 28421725653 residues

Word size : 24
Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Listing first 1000 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	480	100.0	480	6	BD009867
c	2	480	100.0	13222	1 AE000637
3	294	61.3	853	6	AX788995
4	100	20.8	659	6	AX789107
c	5	86	17.9	14404	1 AE001551
6	61	12.7	480	1	AY725427

ALIGNMENTS

RESULT 1
BD009867
LOCUS
DEFINITION
Proteins, in particular membrane proteins, of Helicobacter pylori, their preparation and use.
ACCESSION
BD009867
VERSION
BD009867.1 GI:18638240

480 bp DNA linear PAT 31-JAN-2002

KEYWORDS JP 2001502886-A/8.
SOURCE Helicobacter pylori
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteriaceae; Helicobacter.
REFERENCE 1 (bases 1 to 480)
AUTHORS Knapp,B., Hundt,E. and Schmidt,K.H.
TITLE Proteins, in particular membrane proteins, of Helicobacter pylori, their preparation and use
JOURNAL Patent: JP 2001502886-A 8 06-MAR-2001;
CHIRON BEHRING GMBH & CO
COMMENT OS Helicobacter pylori
PN JP 2001502886-A/8
PD 06-MAR-2001
PF 25-JUL-1997 JP 1998508651
PI 26-JUL-1996 DE 196 30 390.7
PI BERNHARD KNAPP, ERIKA HUNDT, KARL HEINZ SCHMIDT PC
C12N15/31, C07K14/205, C07K16/12, G01N33/53, A61K31/70, A61K39/106, PC
A61K39/395
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..480
FT /organism='Helicobacter pylori'.
FEATURES
source Location/Qualifiers
1..480
/organism='Helicobacter pylori'
/mol_type='genomic DNA'
/db_xref='taxon:210'

Query Match 100.0%; Score 480; DB 6; Length 480;
Best Local Similarity 100.0%; Pred. No. 3.8e-273;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
Db 1 ATGGACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
Qy 61 CTACCTCACCGCTATCCCGTCTTTAGTGATAGATTAATATAGAGTTACAGCCCAATAA 120
Db 61 CTACCTCACCGCTATCCCGTCTTTAGTGATAGATTAATATAGAGTTACAGCCCAATAA 120
Qy 121 AAAATTGTCGCTTATAAGATATACATTTTAAATCAAGACGTTGTTTAAACGGCATTTCCCT 180
Db 121 AAAATTGTCGCTTATAAGATATACATTTTAAATCAAGACGTTGTTTAAACGGCATTTCCCT 180
Qy 181 AATAAGCCCATTTTCCCGCGCTTTTGCATCGTAGAGGCATGCGGCAACCGGAGGTTT 240
Db 181 AATAAGCCCATTTTCCCGCGCTTTTGCATCGTAGAGGCATGCGGCAACCGGAGGTTT 240
Qy 241 TTAGCCTTCACTAGCTTGTGGGGTGTGACCTGAAATCGCCAAACAAAATCGTGAT 300
Db 241 TTAGCCTTCACTAGCTTGTGGGGTGTGACCTGAAATCGCCAAACAAAATCGTGAT 300
Qy 301 TTCAATGACGATTCATAGGTTAAATTCGCGATCCCTGTAAACCCAGCGCAGAGATTAGAA 360
Db 301 TTCAATGACGATTCATAGGTTAAATTCGCGATCCCTGTAAACCCAGCGCAGAGATTAGAA 360
Qy 361 TACCATTTAGAACTCTTAAAGCATAGGGCATGATCTGGCAAGTGGGTGGCACCGCTCAA 420
Db 361 TACCATTTAGAACTCTTAAAGCATAGGGCATGATCTGGCAAGTGGGTGGCACCGCTCAA 420
Qy 421 GTGGATGGCAAGTGGTTCGCTGAAGCCGAATTGAAGCCATGATTGCAGAGAGATTAA 480
Db 421 GTGGATGGCAAGTGGTTCGCTGAAGCCGAATTGAAGCCATGATTGCAGAGAGATTAA 480

RESULT 2
AE000637/c
LOCUS
DEFINITION
Helicobacter pylori 26695 section 115 of 134 of the complete genome.
BCT 06-APR-1999

ACCESSION	AE000637	AE000511
VERSION	AE000637.1	GI:2314536
KEYWORDS		
SOURCE	Helicobacter pylori 26695	
ORGANISM	Helicobacter pylori 26695	
REFERENCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales; Helicobacteraceae; Helicobacter.	
AUTHORS	1 (bases 1 to 13222) Tomb, J.-F., White, O., Kerslavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.	gene CDS
TITLE	The complete genome sequence of the gastric pathogen Helicobacter pylori	
JOURNAL	Nature 388 (6642), 539-547 (1997)	
PUBMED	9252185	
REFERENCE	2 (bases 1 to 13222) Tomb, J.-F., White, O., Kerslavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Cotton, M.D., Weidman, J.M., Fujii, C., Bowman, C., Wathey, L., Wallin, E., Hayes, W.S., Borodovsky, M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.	gene CDS
TITLE	Submitted (06-AUG-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA	
JOURNAL	3 (bases 1 to 13222)	
REFERENCE	White, O.	
AUTHORS	Direct Submission	
TITLE	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA	
JOURNAL	Location/Qualifiers	
FEATURES	1. .13222 /organism="Helicobacter pylori 26695" /mol_type="genomic DNA" /strain="26695" /db_xref="taxon:85962" complement (72. .1343) /gene="HP1366" complement (72. .1343) /gene="HP1366" /note="similar to SP:P23191 GB:X56977 PID:44182 percent identity: 37.14; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="type IIS restriction enzyme R protein (MBOIR)" /protein_id="AAD08410.1" /db_xref="GI:2314537" /translation="MFPLEKLEILLETQDPSKECLFLANIKSSDYRGHLHSQHNRYDQNKITIIQALFNEVGDFLQIRITDMSKRPSNIIIGEEIYAKVDNICKSEMPQDNSGKNQVQDSLRNLFVDMHMGLEIRYKNNKPTNPTQSNIKYISLTPLAIEFLNQLLRKNFCVYQALENLLQFGACREVMIELDNHYLDIIEEMFPVTFILNIENFTREITEYRGYRSLRSIQEKLKELVQDVCNPNHFNGLKRDYHNWKNQAQIIFSLLEQSVFFETNKERLLIKALNEENKQNDKKIKRSIKKALYFEKHGVKKEKGPRLHHI VPLCLARSIEPDLLDKWNELIYDAPNHAKISQTNKHICLIFKNCDDVLSKGLKEQBSLYLTITENLVKLDLQAMUKYNKDLHLSKNG" complement (1359. .2141) /gene="HP1367" complement (1359. .2141) /note="similar to SP:P23192 GB:X56977 PID:44181 percent identity: 59.27; identified by sequence similarity; putative" /codon_start=1 /transl_table=11	gene CDS
gene	/product="type IIS restriction enzyme M1 protein (mod)" /protein_id="AAD08411.1" /db_xref="GI:2314538" /translation="MILNKIYIDVFTFLDKLEDKSDLAIIIDPPYNLKIASWDSFKNDEEFTFSVAWIDKMLPKLGTGSFYIENFPNCALFLAYLHHKPHFLNPIITWVKKOGRANAKKRVNHQAQESILFYSMHKNTFNADERIAVESAEIRIKHAQSGILKNNKRWFPNPKSLDWEITSQRHVEKKGILPKPSPSIKPKALIERMIKASSHKNDLILDLFSGSGMTSLDVAKSLEERNFICGSHAEBYVHGSLEMFRIYBCE" complement (2128. .2991) /gene="HP1368" complement (2128. .2991) /note="HP1368" /note="similar to SP:P29568 percent identity: 32.96; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="type IIS restriction enzyme M2 protein (mod)" /protein_id="AAD08412.1" /db_xref="GI:2314539" /translation="MNINKVYHSSTNMNEVPNSVDLIITSPYFNIKDYAKNGTQDLQHSQHVDELGALERYEDYLLGLKWLCEYKALPKNGKLCINPLMFLMKVYLNTFYNRHI FDLHADIQHSILHDLNNLKNPKMFLDVIYKQANPKTRMLMSGYPYPRNFYAQNTIEFTIGVFVKDGPKOPTREOKQSOLTOSEWVEFTKOIWEIPNKNDLAFGKHAALMPAEALARRILRLSCYGDVLDLPFSGSTTLREAKLLKRNFIGVELYENYKPLIEQKLGMLDFDE" 3240. .4664 /gene="HP1369" /note="This region contains an authentic frame shift and is not the result of a sequencing artifact; similar to SP:P40814 percent identity: 39.74; identified by sequence similarity; putative; type III restriction enzyme M"	gene CDS
gene	/product="type III restriction enzyme M protein (mod)" /protein_id="AAD08413.1" /db_xref="GI:2314540" /translation="MYKLDENGPKLDPVWAKSGNNPSPYTLNGKTWSPSGTFFWRYSIGTLKDMQNNRIVFNGKNPMKRYLSEVAGRKSTPFDGSEVGYNLNGDAEIKQLFNGKNVFNPKPEALLQRILEISTKENDLVDLPFAGSGTTCAVAHKMKRYIGLEMGDEPESVILPRKKVIGGFGKGAAPFDGGVIVKVELESYEETIRKIKYENDRKLAYDEQYSDILVERKNESYTLNIEALEKMGVDIKETLENLWGVGVEFFNEKVKPKGNDKEVEILKALKALIIW" 5679. .8585 /gene="HP1371" 5679. .8585 /gene="HP1371" /note="similar to GB:L25415 PID:496157 percent identity: 26.73; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="type III restriction enzyme R protein" /protein_id="AAD08414.1" /db_xref="GI:2314541" /translation="MAKKKDLITDNEIFVAQKLAEELTNEINEPRLRDLFKSPDNNKEILDYQQALINAFMLVAFRDFPKENKKEFYAQKHYSFAHCDFAKKCLINPLLKSHPKVENHCVSENFINRLAFYMATGSGTIVIKLVELLSVAIRMGILPKKNIMPFSA NENLIQOFEKELEIKYNNRNDYFKQIDFKSLKSVTHKDFYAPKQSVIIQITLIFYRADLMNDEEKENLLANYKYDNGENDYVILDEAHGKNKSBSKQALFSLLSLKGKFNFSATFTESDLITAVYNLSVGWVKGKGLGVKSSVNNKNSDAEIFFKTLARVENDDGDSDLKAKLLLLGMOKRYKTEGYFYDPLMLVFTHSVNNKNSDAEIFFKTLARVENDDGDSDLKAKEDLLEELKNPEFLPSDDDKDYKVKPKKEGKSMDFKGLKEEVFYANNHIEVITNPKNNGEIAFLNTSDKVPGLIRIGDITETWICEKLKSVKVSKNLSFEESYFSQIDKSSINILVGSRTFDTGWDSTPVSILFDLDDDAKLLVKQSPGKGVRIESYKQNRQRLAYLDIDGAIKALKKPNAMLETFLVPTNYTNSLEAILKFQKESNGWREIKLEKTPTKHALFVPCYRKEQTSILPEPSASFPMSEKPNKDLKEYFNLMSEKHFILKHEIYD	gene CDS

```
PKYMLKQKMTQEAHFKNVSTWYKDLDMYMSIKGLKLYPNQKVPKDEFNALDSEKIV
HFKRLKVRADKEELVKTIQEVKEYAPLDKTELIIKKIAOGEIDPYTEKHQKNTKY
GGAELKUKHERYITPLIKRKNCDMLKVVVKVSESDFLRELLKITITLQNDYDFWAPS
FISLSDTHSVSNLPVNLNLEENKIYGLVSHNQAIQIAVLEKGRNLNGFLNAHKRCAYSM
IQONQVLGFIQTNFQELVVDPIVPSAEINIGDQVLTSGLDGIQFGAGVPGVGRVSTED
HYTKSAVLKNKAFSLGAKLLRHVFLSDVKQ"
/sequence similarity; putative"
/codon start=1
/transl_table=11
/product="rod shape-determining protein (mreC)"
/protein id="AAD08415.1"
/db_xref="GI:2314542"
/translation="NRFFYKFLWLLGILFIYFLOPKGSSSYISDRIKNALMNKNSL
LDNQVAFQQAQNIKEFKERLILLEALKLENADLERLNSIYPLENPKMTYPTFTWTS
FISLSDTHSVSNLPVNLNLEENKIYGLVSHNQAIQIAVLEKGRNLNGFLNAHKRCAYSM
IQONQVLGFIQTNFQELVVDPIVPSAEINIGDQVLTSGLDGIQFGAGVPGVGRVSTED
HYTKSAVLKNKAFSLGAKLLRHVFLSDVKQ"
complement(9385..10428)
/gene="HP1373"
complement(9385..10428)
/gene="HP1373"
/note="similar to SP:P13519 GB:M22055 PID:606191
PID:1128967 GB:U00096 percent identity: 51.87; identified
by sequence similarity; putative"
/codon start=1
/transl_table=11
/product="rod shape-determining protein (mreB)"
/protein id="AAD08416.1"
/db_xref="GI:2314543"
/translation="MIFSKLIGLFSHDIAIDLTANTIVLVKGGIILNEPSIVAVRM
GLFDSKADILAVGSEAKMLGKTPNSRAIRPMKDCVIADYDITAKMIRYFIEVKHK
RTWIRPRIMVCVPGLTSVERNAKVBSTLSAGAEVFLIEBPMAAIIAGALPVKPEPQ
GSLIVDIGGTEIGSLGLVIVSKSRVAGDKLDQISIVEYIRKKNLLIGERTGEE
IKIEITGCAIKLDPDPTMEVSGRDQVSGLLHTIELSSDDVFEAIKDQVREISSALRSVL
BEVKPDLAKDIVONGVVLVTGGGALKIGLDKYLSDMVKLPVYVVDPELLAVAKGTGEAI
QDLDLLSRVGRSE"
complement(10471..11811)
/gene="HP1374"
complement(10471..11811)
/gene="HP1374"
/note="similar to GB:JL18867 SP:P33138 GB:223278 PID:347729
PID:388977 percent identity: 55.04; identified by sequence
similarity; putative"
/codon start=1
/transl_table=11
/product="ATP-dependent protease ATPase subunit (clpX)"
Query Match 100.0%; Score 480; DB 1; Length 13222;
Best Local Similarity 100.0%; Pred. No. 3.6e-273;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAACAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
DB 13108 ATGGAACAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 13049
QY 61 CTACCTCACCGCTATCCCATCTTATAGATAGAAATTTATAGAGTTACAAGCCCAATAA 120
DB 13048 CTACCTCACCGCTATCCCATCTTATAGATAGAAATTTATAGAGTTACAAGCCCAATAA 12989
QY 121 AAAATTGCGCTTATAGAAATATCACTTTTAAAGACGCTGTTTAAACCGGCAATTCCTT 180
DB 12988 AAAATTGCGCTTATAGAAATATCACTTTTAAAGACGCTGTTTAAACCGGCAATTCCTT 12929
QY 181 AATAAGCCATTTTCCCGGCGTTTGTATCGTAGAGGCATCGCCAAACGGAGGGTTT 240
DB 12928 AATAAGCCATTTTCCCGGCGTTTGTATCGTAGAGGCATCGCCAAACGGAGGGTTT 12869
QY 241 TTAGCCTTCACTAGCTTGTGGGGGTTTGACCTCGAAATCGCCAAACAAAAATCGTGAT 300
```

```
DB 12868 TTAGCCTTCACTAGCTTGTGGGGGTTTGACCTCGAAATCGCCAAACAAAAATCGTGAT 12809
QY 301 TTATGATACGATTGATAAGGTTAAATTCGCGATCCCTGTAAACCCAGCGACAGATTAGAA 360
DB 12808 TTATGATACGATTGATAAGGTTAAATTCGCGATCCCTGTAAACCCAGCGACAGATTAGAA 12749
QY 361 TACCATTTAGAACTCTTAAAGCATAAGGGCATGATCGGCAAGTGGGTGGCACGGCTCAA 420
DB 12748 TACCATTTAGAACTCTTAAAGCATAAGGGCATGATCGGCAAGTGGGTGGCACGGCTCAA 12689
QY 421 GTGGATCGCAAAAGTGGTTCGCTGAAGCCGAAATTGAAAAGCCATGATTGACAGAGAGATTAA 480
DB 12688 GTGGATCGCAAAAGTGGTTCGCTGAAGCCGAAATTGAAAAGCCATGATTGACAGAGAGATTAA 12629
RESULT 3
AX788995 853 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Sequence 1459 from Patent WO02066501.
ACCESSION AX788995
VERSION AX788995.1 GI:32955341
KEYWORDS
SOURCE Helicobacter pylori
ORGANISM Helicobacter pylori
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1
AUTHORS Legrain, P., Rain, J. C., Colland, F., de Reuse, H. and Labigne, A.
TITLE Protein-protein interactions in Helicobacter pylori
JOURNAL Patent: WO 02066501-A 1459 29-AUG-2002;
Hybrigenics (FR) ; INSTITUT PASTEUR (FR)
FEATURES
source 1..853
/organism="Helicobacter pylori"
/mol_type="unassigned DNA"
/db_xref="taxon:210"
ORIGIN
Query Match 61.3%; Score 294; DB 6; Length 853;
Best Local Similarity 100.0%; Pred. No. 1.4e-162;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAACAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
DB 560 ATGGAACAAGCCATCAAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 619
QY 61 CTACCTCACCGCTATCCCATCTTATAGTGAATAGAAATTTATAGAGTTACAAGCCCAATAA 120
DB 620 CTACCTCACCGCTATCCCATCTTATAGTGAATAGAAATTTATAGAGTTACAAGCCCAATAA 679
QY 121 AAAATTGCGCTTATAGAAATATCACTTTTAAAGACGCTGTTTAAAGCGGCAATTCCTT 180
DB 680 AAAATTGCGCTTATAGAAATATCACTTTTAAAGACGCTGTTTAAAGCGGCAATTCCTT 739
QY 181 AATAAGCCATTTTCCCGGCGTTTGTATCGTAGAGGCATGGCCAAACGGAGGGTTT 240
DB 740 AATAAGCCATTTTCCCGGCGTTTGTATCGTAGAGGCATGGCCAAACGGAGGGTTT 799
QY 241 TTAGCCTTCACTAGCTTGTGGGGGTTTGACCTCGAAATCGCCAAACAAAAATC 294
DB 800 TTAGCCTTCACTAGCTTGTGGGGGTTTGACCTCGAAATCGCCAAACAAAAATC 853
RESULT 4
AX789107 659 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Sequence 1571 from Patent WO02066501.
ACCESSION AX789107
VERSION AX789107.1 GI:32955397
KEYWORDS
SOURCE Helicobacter pylori
ORGANISM Helicobacter pylori
```

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.

REFERENCE

AUTHORS Legrain, P., Rain, J.C., Colland, F., de Reuse, H. and Labigne, A.
TITLE Protein-protein interactions in Helicobacter pylori
JOURNAL Patent: WO 02066501-A 1571 29-AUG-2002;
Hybrigenics (FR) ; INSTITUT PASTEUR (FR)

FEATURES

source
1..659
/organism="Helicobacter pylori"
/mol_type="unassigned DNA"
/db_xref="taxon:210"

ORIGIN

Query Match 20.8%; Score 100; DB 6; Length 659;
Best Local Similarity 100.0%; Pred. No. 3.1e-47;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
|||||
Db 560 ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 619
|||||
QY 61 CTACCTCACCGTATCCATGCTTTTAGTGGATAGAAATTA 100
|||||
Db 620 CTACCTCACCGTATCCATGCTTTTAGTGGATAGAAATTA 659
|||||

RESULT 5

AB001551/c

LOCUS

DEFINITION Helicobacter pylori, strain J99 section 112 of 132 of the complete genome.

ACCESSION

AB001551 AB001439

VERSION

AB001551.1 GI:4155887

KEYWORDS

SOURCE

ORGANISM

Helicobacter pylori J99

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.

REFERENCE

AUTHORS

Alm, R.A., Ling, L.S., Moir, D.T., King, B.L., Brown, E.D., Doig, P.C.,
Smith, D.R., Noonan, B., Guild, B.C., deJonge, B.L., Carmel, G.,
Tummino, P.J., Caruso, A., Uria-Nickelsen, M., Mills, D.M., Ives, C.,
Gibson, R., Merberg, D., Mills, S.D., Jiang, Q., Taylor, D.E.,
Vovis, G.F. and Trust, T.J.
Genomic-sequence comparison of two unrelated isolates of the human
gastric pathogen Helicobacter pylori
Nature 397 (6715), 176-180 (1999)

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

King, B.L., Alm, R.A. and Trust, T.J.
Direct Submission
Submitted (12-JAN-1999) Astra Research Center Boston, 128 Sidney
Street, Cambridge, MA 02139, USA
Address all correspondence to: hpoarch.us.astra.com or Richard
A. Alm, Astra Research Center Boston, 128 Sidney Street, Cambridge,
MA, 02139. Lo-Seel, L. Ling, Donald T. Moir, Douglas R. Smith,
Braydon C. Guild, Gilles Carmel, Anthony Caruso, Debra M. Mills,
Rene Gibson, and Gerald F. Vovis are with Genome Therapeutics
Corporation, 100 Beaver Street, Waltham, MA, 02453. Qin Jiang and
Diane E. Taylor are with the University of Alberta Department of
Medical Microbiology and Immunology, Edmonton, Alberta, Canada. T6G
2H7 and the Canadian Bacterial Diseases Network. All other authors
are with Astra Research Center Boston, 128 Sidney Street,
Cambridge, MA, 02139. Putative identifications, sequence
alignments, and name and sequence search capability are available
at ARCB's World Wide web site. (URL:
http://www.astra-boston.com/hpylori).

FEATURES

source

1..14404
/organism="Helicobacter pylori J99"
/mol_type="genomic DNA"
/strain="J99"

gene

CDS

/db_xref="taxon:85963"
complement (71..1531)
/gene="dnaB"
/note="synonym: jhp1280"
complement (71..1531)
/gene="dnaB"
/note="similar to H. pylori 26695 gene HP1362"
/codon_start=1
/transl_table=11
/product="REPLICATIVE DNA HELICASE"
/protein_id="AAD06865.1"
/db_xref="GI:4155895"
/translation="MDHLKHLQQLQNIERIVLSGIVLANHKIIEIHVSLEPSPDYFP
HGLFFELAKLHEVNCPIDENFRQKPKDKQISEDDLVAIFAAASPIDNIEAYVEEIK
NASIKRKLFLANTIREQALSAQKSDIILNAVEREYVALLNGSTIEGFGIKVEVLES
TNWLTENQKGLSKVTGPTGVQLDNVTSGFNQSLVILGARSPMGKTSLMNMVVL
SALNDRGVAVFSLEMSAEQLARLSDLTSLNNHDLSEARLDDOWENLAKCFDHL
QKKLFYDKSYVRMDQIRLQKLSQHKELGIAFDYDLQMSGNKATKEHEQIAEI
SRELKTLARELEIPIIALVOLNKSLENRRDKPILSDIKOSGGIEQADIVLFYRGY
IYOMRAEDNKIDKLKKEGVBEAQELHLKYNEERRIHKQNGSTIEEABIIIVAKNRNGAT
GTVYTRFNAPFTRYEDPVDVSHLEGGQETKFEMPTT"
complement (1542..2936)
/gene="jhp1281"
complement (1542..2936)
/gene="jhp1281"
/note="similar to H. pylori 26695 gene HP1363"
/codon_start=1
/transl_table=11
/product="putative"
/protein_id="AAD06866.1"
/db_xref="GI:4155896"
/translation="MLSVEYKGNALDKRVLEWLLSEDIIMENAAALERAVLQNASL
GAKVIILCGSDNGDGYTLARLVGRFPTLVFEMKLAKSPMCOLOKRAKKGVVVK
AWKEKNEDECVDVDCVWGSARFGLPEFLPESLSQAKARFKIACDIPSGIDSKGRV
DKFAFKADTTISMGALSKCLISDKADYIGELKVGHGVNQIYVEIPTDTLLKSDLL
KLPRDRKNAHKDYGHAHVLLGKHSAGLISALSALSPGSGVGVSIQALECEITSNNK
PLIEVFCENFPKLSFALMGLENIPDPFKKMLEAPCVLDGAGVFTKVEQLALEKE
VILTPHKEFLSLIKSGINISMLELDNKLLETARDFSQKYPKVLLILKGTANTIAHQ
GRVFINLGSVALAKSGDVLGVLISLLSQNTYPLDAAINASLAHALAGLEFKHY
ALTPDLQLEIKREL"
complement (2939..3973)
/gene="jhp1282"
complement (2939..3973)
/gene="jhp1282"
/note="similar to H. pylori 26695 gene HP1364"
/codon_start=1
/transl_table=11
/product="putative histidine kinase sensor protein"
/protein_id="AAD06867.1"
/db_xref="GI:4155897"
/translation="MASHKIASEIIQLHMOTHADYHNALNALSRYKDASIALFDSKKR
VLYSNIPESADLIKTHKEGFSFGKGVYLFSDTEAHLGVAKMLFKNSKPLHPSLV
RNWLVFVAFVLCVIGVSVFLGRILFKPIRNEIIRIDHFLXNTTTHLTPSALVSL
KTLDEQQHRRIKIAIQORMSFYRSLSYLVQMDIERESFALLDLKALTENTLFPSEM
IDYHLEKSDLVGVVLKAEQDFISLYSNLNNAIKYSVMHGYIHLTLPFLKVK
LGVEIPDKITKLSVRYAREFNSSVLGYGIGLDLVKKVCEKYKMRLEIHSSEPSKGFH
ENSFCVQFOQ"
complement (4107..4748)
/note="similar to H. pylori 26695 gene HP1365"
/codon_start=1
/transl_table=11
/product="putative TRANSCRIPTIONAL REGULATOR"
/protein_id="AAD06868.1"
/db_xref="GI:4165002"
/translation="MQKKIFLEDYLLSISIKFLEHLYGEVFCAFNGKAEHRLSV
ERNLLLDVQVPEMSLELFKIKNDPLISTPVIPTALQDNATLKNALNGASDYL
KKPFDLDEARIKRFENDDPIEIMPNI FYVHALNVKKEEIIAPKTAQLLEYLH
KGQIISQALENNLWEQAIDDTLRTYIKVLRLKLGKNCIETHKGVGYRFPNL"
complement (4110..4748)
/gene="jhp1283"
4857..6806
/gene="jhp1284"
4857..6806

gene

CDS

/db_xref="taxon:85963"
complement (71..1531)
/gene="dnaB"
/note="synonym: jhp1280"
complement (71..1531)
/gene="dnaB"
/note="similar to H. pylori 26695 gene HP1362"
/codon_start=1
/transl_table=11
/product="REPLICATIVE DNA HELICASE"
/protein_id="AAD06865.1"
/db_xref="GI:4155895"
/translation="MDHLKHLQQLQNIERIVLSGIVLANHKIIEIHVSLEPSPDYFP
HGLFFELAKLHEVNCPIDENFRQKPKDKQISEDDLVAIFAAASPIDNIEAYVEEIK
NASIKRKLFLANTIREQALSAQKSDIILNAVEREYVALLNGSTIEGFGIKVEVLES
TNWLTENQKGLSKVTGPTGVQLDNVTSGFNQSLVILGARSPMGKTSLMNMVVL
SALNDRGVAVFSLEMSAEQLARLSDLTSLNNHDLSEARLDDOWENLAKCFDHL
QKKLFYDKSYVRMDQIRLQKLSQHKELGIAFDYDLQMSGNKATKEHEQIAEI
SRELKTLARELEIPIIALVOLNKSLENRRDKPILSDIKOSGGIEQADIVLFYRGY
IYOMRAEDNKIDKLKKEGVBEAQELHLKYNEERRIHKQNGSTIEEABIIIVAKNRNGAT
GTVYTRFNAPFTRYEDPVDVSHLEGGQETKFEMPTT"
complement (1542..2936)
/gene="jhp1281"
complement (1542..2936)
/gene="jhp1281"
/note="similar to H. pylori 26695 gene HP1363"
/codon_start=1
/transl_table=11
/product="putative"
/protein_id="AAD06866.1"
/db_xref="GI:4155896"
/translation="MLSVEYKGNALDKRVLEWLLSEDIIMENAAALERAVLQNASL
GAKVIILCGSDNGDGYTLARLVGRFPTLVFEMKLAKSPMCOLOKRAKKGVVVK
AWKEKNEDECVDVDCVWGSARFGLPEFLPESLSQAKARFKIACDIPSGIDSKGRV
DKFAFKADTTISMGALSKCLISDKADYIGELKVGHGVNQIYVEIPTDTLLKSDLL
KLPRDRKNAHKDYGHAHVLLGKHSAGLISALSALSPGSGVGVSIQALECEITSNNK
PLIEVFCENFPKLSFALMGLENIPDPFKKMLEAPCVLDGAGVFTKVEQLALEKE
VILTPHKEFLSLIKSGINISMLELDNKLLETARDFSQKYPKVLLILKGTANTIAHQ
GRVFINLGSVALAKSGDVLGVLISLLSQNTYPLDAAINASLAHALAGLEFKHY
ALTPDLQLEIKREL"
complement (2939..3973)
/gene="jhp1282"
complement (2939..3973)
/gene="jhp1282"
/note="similar to H. pylori 26695 gene HP1364"
/codon_start=1
/transl_table=11
/product="putative histidine kinase sensor protein"
/protein_id="AAD06867.1"
/db_xref="GI:4155897"
/translation="MASHKIASEIIQLHMOTHADYHNALNALSRYKDASIALFDSKKR
VLYSNIPESADLIKTHKEGFSFGKGVYLFSDTEAHLGVAKMLFKNSKPLHPSLV
RNWLVFVAFVLCVIGVSVFLGRILFKPIRNEIIRIDHFLXNTTTHLTPSALVSL
KTLDEQQHRRIKIAIQORMSFYRSLSYLVQMDIERESFALLDLKALTENTLFPSEM
IDYHLEKSDLVGVVLKAEQDFISLYSNLNNAIKYSVMHGYIHLTLPFLKVK
LGVEIPDKITKLSVRYAREFNSSVLGYGIGLDLVKKVCEKYKMRLEIHSSEPSKGFH
ENSFCVQFOQ"
complement (4107..4748)
/note="similar to H. pylori 26695 gene HP1365"
/codon_start=1
/transl_table=11
/product="putative TRANSCRIPTIONAL REGULATOR"
/protein_id="AAD06868.1"
/db_xref="GI:4165002"
/translation="MQKKIFLEDYLLSISIKFLEHLYGEVFCAFNGKAEHRLSV
ERNLLLDVQVPEMSLELFKIKNDPLISTPVIPTALQDNATLKNALNGASDYL
KKPFDLDEARIKRFENDDPIEIMPNI FYVHALNVKKEEIIAPKTAQLLEYLH
KGQIISQALENNLWEQAIDDTLRTYIKVLRLKLGKNCIETHKGVGYRFPNL"
complement (4110..4748)
/gene="jhp1283"
4857..6806
/gene="jhp1284"
4857..6806

gene

CDS

/db_xref="taxon:85963"
complement (71..1531)
/gene="dnaB"
/note="synonym: jhp1280"
complement (71..1531)
/gene="dnaB"
/note="similar to H. pylori 26695 gene HP1362"
/codon_start=1
/transl_table=11
/product="REPLICATIVE DNA HELICASE"
/protein_id="AAD06865.1"
/db_xref="GI:4155895"
/translation="MDHLKHLQQLQNIERIVLSGIVLANHKIIEIHVSLEPSPDYFP
HGLFFELAKLHEVNCPIDENFRQKPKDKQISEDDLVAIFAAASPIDNIEAYVEEIK
NASIKRKLFLANTIREQALSAQKSDIILNAVEREYVALLNGSTIEGFGIKVEVLES
TNWLTENQKGLSKVTGPTGVQLDNVTSGFNQSLVILGARSPMGKTSLMNMVVL
SALNDRGVAVFSLEMSAEQLARLSDLTSLNNHDLSEARLDDOWENLAKCFDHL
QKKLFYDKSYVRMDQIRLQKLSQHKELGIAFDYDLQMSGNKATKEHEQIAEI
SRELKTLARELEIPIIALVOLNKSLENRRDKPILSDIKOSGGIEQADIVLFYRGY
IYOMRAEDNKIDKLKKEGVBEAQELHLKYNEERRIHKQNGSTIEEABIIIVAKNRNGAT
GTVYTRFNAPFTRYEDPVDVSHLEGGQETKFEMPTT"
complement (1542..2936)
/gene="jhp1281"
complement (1542..2936)
/gene="jhp1281"
/note="similar to H. pylori 26695 gene HP1363"
/codon_start=1
/transl_table=11
/product="putative"
/protein_id="AAD06866.1"
/db_xref="GI:4155896"
/translation="MLSVEYKGNALDKRVLEWLLSEDIIMENAAALERAVLQNASL
GAKVIILCGSDNGDGYTLARLVGRFPTLVFEMKLAKSPMCOLOKRAKKGVVVK
AWKEKNEDECVDVDCVWGSARFGLPEFLPESLSQAKARFKIACDIPSGIDSKGRV
DKFAFKADTTISMGALSKCLISDKADYIGELKVGHGVNQIYVEIPTDTLLKSDLL
KLPRDRKNAHKDYGHAHVLLGKHSAGLISALSALSPGSGVGVSIQALECEITSNNK
PLIEVFCENFPKLSFALMGLENIPDPFKKMLEAPCVLDGAGVFTKVEQLALEKE
VILTPHKEFLSLIKSGINISMLELDNKLLETARDFSQKYPKVLLILKGTANTIAHQ
GRVFINLGSVALAKSGDVLGVLISLLSQNTYPLDAAINASLAHALAGLEFKHY
ALTPDLQLEIKREL"
complement (2939..3973)
/gene="jhp1282"
complement (2939..3973)
/gene="jhp1282"
/note="similar to H. pylori 26695 gene HP1364"
/codon_start=1
/transl_table=11
/product="putative histidine kinase sensor protein"
/protein_id="AAD06867.1"
/db_xref="GI:4155897"
/translation="MASHKIASEIIQLHMOTHADYHNALNALSRYKDASIALFDSKKR
VLYSNIPESADLIKTHKEGFSFGKGVYLFSDTEAHLGVAKMLFKNSKPLHPSLV
RNWLVFVAFVLCVIGVSVFLGRILFKPIRNEIIRIDHFLXNTTTHLTPSALVSL
KTLDEQQHRRIKIAIQORMSFYRSLSYLVQMDIERESFALLDLKALTENTLFPSEM
IDYHLEKSDLVGVVLKAEQDFISLYSNLNNAIKYSVMHGYIHLTLPFLKVK
LGVEIPDKITKLSVRYAREFNSSVLGYGIGLDLVKKVCEKYKMRLEIHSSEPSKGFH
ENSFCVQFOQ"
complement (4107..4748)
/note="similar to H. pylori 26695 gene HP1365"
/codon_start=1
/transl_table=11
/product="putative TRANSCRIPTIONAL REGULATOR"
/protein_id="AAD06868.1"
/db_xref="GI:4165002"
/translation="MQKKIFLEDYLLSISIKFLEHLYGEVFCAFNGKAEHRLSV
ERNLLLDVQVPEMSLELFKIKNDPLISTPVIPTALQDNATLKNALNGASDYL
KKPFDLDEARIKRFENDDPIEIMPNI FYVHALNVKKEEIIAPKTAQLLEYLH
KGQIISQALENNLWEQAIDDTLRTYIKVLRLKLGKNCIETHKGVGYRFPNL"
complement (4110..4748)
/gene="jhp1283"
4857..6806
/gene="jhp1284"
4857..6806

gene

CDS

/db_xref="taxon:85963"
complement (71..1531)
/gene="dnaB"
/note="synonym: jhp1280"
complement (71..1531)
/gene="dnaB"
/note="similar to H. pylori 26695 gene HP1362"
/codon_start=1
/transl_table=11
/product="REPLICATIVE DNA HELICASE"
/protein_id="AAD06865.1"
/db_xref="GI:4155895"
/translation="MDHLKHLQQLQNIERIVLSGIVLANHKIIEIHVSLEPSPDYFP
HGLFFELAKLHEVNCPIDENFRQKPKDKQISEDDLVAIFAAASPIDNIEAYVEEIK
NASIKRKLFLANTIREQALSAQKSDIILNAVEREYVALLNGSTIEGFGIKVEVLES
TNWLTENQKGLSKVTGPTGVQLDNVTSGFNQSLVILGARSPMGKTSLMNMVVL
SALNDRGVAVFSLEMSAEQLARLSDLTSLNNHDLSEARLDDOWENLAKCFDHL
QKKLFYDKSYVRMDQIRLQKLSQHKELGIAFDYDLQMSGNKATKEHEQIAEI
SRELKTLARELEIPIIALVOLNKSLENRRDKPILSDIKOSGGIEQADIVLFYRGY
IYOMRAEDNKIDKLKKEGVBEAQELHLKYNEERRIHKQNGSTIEEABIIIVAKNRNGAT
GTVYTRFNAPFTRYEDPVDVSHLEGGQETKFEMPTT"
complement (1542..2936)
/gene="jhp1281"
complement (1542..2936)
/gene="jhp1281"
/note="similar to H. pylori 26695 gene HP1363"
/codon_start=1
/transl_table=11
/product="putative"
/protein_id="AAD06866.1"
/db_xref="GI:4155896"
/translation="MLSVEYKGNALDKRVLEWLLSEDIIMENAAALERAVLQNASL
GAKVIILCGSDNGDGYTLARLVGRFPTLVFEMKLAKSPMCOLOKRAKKGVVVK
AWKEKNEDECVDVDCVWGSARFGLPEFLPESLSQAKARFKIACDIPSGIDSKGRV
DKFAFKADTTISMGALSKCLISDKADYIGELKVGHGVNQIYVEIPTDTLLKSDLL
KLPRDRKNAHKDYGHAHVLLGKHSAGLISALSALSPGSGVGVSIQALECEITSNNK
PLIEVFCENFPKLSFALMGLENIPDPFKKMLEAPCVLDGAGVFTKVEQLALEKE
VILTPHKEFLSLIKSGINISMLELDNKLLETARDFSQKYPKVLLILKGTANTIAHQ
GRVFINLGSVALAKSGDVLGVLISLLSQNTYPLDAAINASLAHALAGLEFKHY
ALTPDLQLEIKREL"
complement (2939..3973)
/gene="jhp1282"
complement (2939..3973)
/gene="jhp1282"
/note="similar to H. pylori 26695 gene HP1364"
/codon_start=1
/transl_table=11
/product="putative histidine kinase sensor protein"
/protein_id="AAD06867.1"
/db_xref="GI:4155897"
/translation="MASHKIASEIIQLHMOTHADYHNALNALSRYKDASIALFDSKKR
VLYSNIPESADLIKTHKEGFSFGKGVYLFSDTEAHLGVAKMLFKNSKPLHPSLV
RNWLVFVAFVLCVIGVSVFLGRILFKPIRNEIIRIDHFLXNTTTHLTPSALVSL
KTLDEQQHRRIKIAIQORMSFYRSLSYLVQMDIERESFALLDLKALTENTLFPSEM
IDYHLEKSDLVGVVLKAEQDFISLYSNLNNAIKYSVMHGYIHLTLPFLKVK
LGVEIPDKITKLSVRYAREFNSSVLGYGIGLDLVKKVCEKYKMRLEIHSSEPSKGFH
ENSFCVQFOQ"
complement (4107..4748)
/note="similar to H. pylori 26695 gene HP1365"
/codon_start=1
/transl_table=11
/product="putative TRANSCRIPTIONAL REGULATOR"
/protein_id="AAD06868.1"
/db_xref="GI:4165002"
/translation="MQKKIFLEDYLLSISIKFLEHLYGEVFCAFNGKAEHRLSV
ERNLLLDVQVPEMSLELFKIKNDPLISTPVIPTALQDNATLKNALNGASDYL
KKPFDLDEARIKRFENDDPIEIMPNI FYVHALNVKKEEIIAPKTAQLLEYLH
KGQIISQALENNLWEQAIDDTLRTYIKVLRLKLGKNCIETHKGVGYRFPNL"
complement (4110..4748)
/gene="jhp1283"
4857..6806
/gene="jhp1284"
4857..6806

gene

gene

CDS

CDS


```

/translation="WIFSKLGLFSDIAIDLGTANTIVLVKGGIILNEPSIVAVRM
GLFDSKAYDILAVGSEAKEMLGKTPNSIRAIRPMKDVIAVDYDIATAKIRYIEKAHK
RTWIRPRIMVCPYIGLTVSVERNAVKSALRSAGAREVFLIEPEMAAIGAGLPVKEPQ
GSLIVDIGGGTTEIGVLSGLVIVSKIRVAGDKLDQSIIVYIRKFNLLIGERTGEE
IKIBIGCAIKLDPPLTMEVSRDQVSGLLHTIELSSDDVFEAIKQDVREISSALRSVL
BEVKPDLVRDIVQGVVLTGGGALIKGLDKVLSDMVKLPVYVGDPELLAVAKGTGEAI
QDLDLISRVGFS"
complement(11630..12988)
/feature="clpX"
/feature="synonym:jhp1288"

gene
Query Match 17.9%; Score 86; DB 1; Length 14404;
Best Local Similarity 100.0%; Pred. No. 6.3e-39;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 GTGTTTAAAGCGGCATTTCCCTAATAAGCCCATTTTCCCGGGCGTTTTCATCGTAGAGGC 219
Db 14125 GTGTTTAAAGCGGCATTTCCCTAATAAGCCCATTTTCCCGGGCGTTTTCATCGTAGAGGC 14066

Qy 220 ATGGCGCAACGGGAGGTTTTCAGC 245
Db 14065 ATGGCGCAACGGGAGGTTTTCAGC 14040

RESULT 6
AY725427 480 bp DNA linear BCT 08-JUL-2005
LOCUS Helicobacter pylori (3R)-hydroxymyristoyl-acyl carrier protein
DEFINITION dehydratase (fabZ) gene, complete cds.
ACCESSION AY725427
VERSION AY725427.1 GI:56684724
KEYWORDS Helicobacter pylori
SOURCE Helicobacter pylori
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteriaceae; Helicobacter.
REFERENCE 1 (bases 1 to 480)
AUTHORS Liu, W., Luo, C., Han, C., Peng, S., Yang, Y., Yue, J., Shen, X. and
Jiang, H.
TITLE A new beta-hydroxyacyl-acyl carrier protein dehydratase (fabZ) from
Helicobacter pylori: Molecular cloning, enzymatic characterization,
and structural modeling
JOURNAL Biochem. Biophys. Res. Commun. 333 (4), 1078-1086 (2005)
PUBMED 15967411
REFERENCE 2 (bases 1 to 480)
AUTHORS Liu, W., Han, C., Wang, Q., Jiang, H. and Shen, X.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-2004) Drug Discovery and Design Center, Shanghai
Institute of Materia Medica, Shanghai Institutes for Biological
Sciences, Chinese Academy of Sciences, Zuchongzhi Road 555,
Shanghai 201203, China
FEATURES
source
Location/Qualifiers
1..480
/organism="Helicobacter pylori"
/mol_type="genomic DNA"
/strain="SS1"
/db_xref="taxon:210"
1..480
/genes="fabZ"
/feature="fabZ"
/function="fatty acid and phosphatidic acid biosynthesis"
/note="enzyme"
/codon_start=1
/transl_table=11
/product="(3R)-hydroxymyristoyl-acyl carrier protein
dehydratase"
/db_xref="GI:56684725"
/db_xref="AAW22049.1"

gene
CDS
/feature="synonym:jhp1287"
/feature="synonym:(11534..11577)
/feature="mreB"
/feature="similar to H. pylori 26695 gene HP1373"
/codon_start=1
/transl_table=11
/product="ROD SHAPE-DETERMINING PROTEIN"
/protein_id="AAD06861.1"
/db_xref="GI:4155891"

gene
CDS
complement(9784..10530)
/feature="jhp1286"
complement(9784..10530)
/feature="jhp1286"
/feature="similar to H. pylori 26695 gene HP1372"
/codon_start=1
/transl_table=11
/product="putative"
/protein_id="AAD06860.1"
/db_xref="GI:4155890"
/translation="MRFFKFLWLGILFIYFLDIKGSYSIDRVKSLMSAKNSL
LNVQAYFQQAQNIKEFKERLILRALKLENADLKERLSITPLENPMTYPTFTNIS
FINLEDTHSVLSNPLNLEENKIYGLVHNOAIGIAVLEKGRGLNLAHKRCAYSM
IQONQVLGFTGNFKQGLVDFIVFVSAEINIQQVLTSLGLDIFGAGVFGVGVSVED
HYTKSAVLKNAPLSEAKLRHVFLSGVKN"
complement(10534..11577)
/feature="mreB"
/feature="synonym:jhp1287"
complement(11534..11577)
/feature="mreB"
/feature="similar to H. pylori 26695 gene HP1373"
/codon_start=1
/transl_table=11
/product="ROD SHAPE-DETERMINING PROTEIN"
/protein_id="AAD06861.1"
/db_xref="GI:4155891"

/translation="MSESTYKDCPCPELIGLNFEDTSLLESKYNEKVDLIDPKLSI
VKDKTKYTNLNSVLTSSKGTQTKNEILKECSNQGVKLDDEILKAFVFKKQGSN
DYFIHNAQGLKEQDLIDLYLFYLFKEMTAFDHRKRLNGINTIKEVALEVLALVSEFEN
ELCKIWNKRLVLSHFIVSLDKAKRYDLSKITNHPYPKQVKEWQDILNLEIADNL
LGNEFLPDLTYFKLEEEVKSLEFNEINGTLIKSENYOALSKNRYKEALDCIYI
DPPNTGSPAYIDRFQDSFWLSLMENRLQALVDLSPOGNFYHLHLDYRANLYGRMLL
NDLFSKERNFNEIIMHRTYQGOIQSNFPRKHSLLWYSCNKNVNFPKITYSDNYKD
TVDRWRWFIVDNNKIYVNPYKADSRFDGLYKLYQSKYKPNNGDIATINGYVD
DWDTDIOADDPKADRELQTLTKQPEKLERIKASSNENSVICDPFAGSGTTCAVA
HKLKRYIGVMEGESEVILPRKPKVIGFGKSCALKBFNGGVIKYTELESEYIEILR
KIKYENDKPLAYEEQVSDLVBRKNSYTLNLEALENMGVDIKETLENLHGVGVFEPN
EKVYKFGNDKEVSLKALKEALI"
6816..9734
/feature="jhp1285"
6816..9734
/feature="jhp1285"
/feature="similar to H. pylori 26695 gene HP1371"
/codon_start=1
/transl_table=11
/product="putative"
/protein_id="AAD06859.1"
/db_xref="GI:4155889"
/translation="MAKKQEVNNEIFVAQKLABEELNANEINDPLEMLDFKSPDN
KELLDYQQALINAFVLVAYFRDPFKSEKFEYAFQKHYSFANCPAKKMLNHLKS
HFKVNHCVRFENIRLAFYATGSGKTIIVIKVELLSVAMGMGLIPKNNMFFSA
NHLIKQFEKIEKYNKDYKQIDFNKLSKNDKYRPAKDSLMEKLETALFYFAD
LMSDESKNLNLYKDCWNGENYVILDBAHKNGKTESKROAIFSLISLKGFLPNFSA
TTBESDLTAVNLSVGEWKLGVGKSVLLKKNLNAFKELKOLDNEKBTALLKA
LLLLGNQKRYEGVGPDLMLVFTVSHVNMNSDAEIPFKTLARVINDESDPSKAK
DDBLEELKPEFLSDGKDKDYKIEVFKESLGDMDFKGLKEAVFYASNGHLEVIIN
PKNQBIAFPAKNTSDVKFCLIRIGDITETWIREKLKSVKVSKNLSFKESYFSQIDKS
SNILVGSRAFTQDWDSPSVILFINLGLDDDAKLVKQSGFGRVIRIESVKNQRRL
AYLEIDEALCKLKPNAEMLEMFVPTNHASLEATLKFKGSENGENRGPWREIKL
EXTRLEHLPVPCYKEQTNALKISQASFKMSKFNKDKLKHFLHMSKHFILKHEI
YDPKYALLKEMIQTAHFKKVSTWHYKOLDYMISEIKGLYPNQKVPKDFEPNALNEK
IYHFKRVKVKADKEKLIQIQEVKEHAPLDKETLRKLAQGEIDPIDTEKKHQDRTP
FKVIDELHNLPTIDYDNATERFPDFIFWLQKGTQIICFIDPKGSKHTDYEHKAD
AYQLPEDKFNPKDDPNLKIYVLFYGDKDDVGERYDLWIEKGLKLEYFFSGFKGLA
"
complement(9784..10530)
/feature="jhp1286"
complement(9784..10530)
/feature="jhp1286"
/feature="similar to H. pylori 26695 gene HP1372"
/codon_start=1
/transl_table=11
/product="putative"
/protein_id="AAD06860.1"
/db_xref="GI:4155890"
/translation="MRFFKFLWLGILFIYFLDIKGSYSIDRVKSLMSAKNSL
LNVQAYFQQAQNIKEFKERLILRALKLENADLKERLSITPLENPMTYPTFTNIS
FINLEDTHSVLSNPLNLEENKIYGLVHNOAIGIAVLEKGRGLNLAHKRCAYSM
IQONQVLGFTGNFKQGLVDFIVFVSAEINIQQVLTSLGLDIFGAGVFGVGVSVED
HYTKSAVLKNAPLSEAKLRHVFLSGVKN"
complement(10534..11577)
/feature="mreB"
/feature="synonym:jhp1287"
complement(11534..11577)
/feature="mreB"
/feature="similar to H. pylori 26695 gene HP1373"
/codon_start=1
/transl_table=11
/product="ROD SHAPE-DETERMINING PROTEIN"
/protein_id="AAD06861.1"
/db_xref="GI:4155891"

/translation="WIFSKLGLFSDIAIDLGTANTIVLVKGGIILNEPSIVAVRM
GLFDSKAYDILAVGSEAKEMLGKTPNSIRAIRPMKDVIAVDYDIATAKIRYIEKAHK
RTWIRPRIMVCPYIGLTVSVERNAVKSALRSAGAREVFLIEPEMAAIGAGLPVKEPQ
GSLIVDIGGGTTEIGVLSGLVIVSKIRVAGDKLDQSIIVYIRKFNLLIGERTGEE
IKIBIGCAIKLDPPLTMEVSRDQVSGLLHTIELSSDDVFEAIKQDVREISSALRSVL
BEVKPDLVRDIVQGVVLTGGGALIKGLDKVLSDMVKLPVYVGDPELLAVAKGTGEAI
QDLDLISRVGFS"
complement(11630..12988)
/feature="clpX"
/feature="synonym:jhp1288"

gene
Query Match 17.9%; Score 86; DB 1; Length 14404;
Best Local Similarity 100.0%; Pred. No. 6.3e-39;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 GTGTTTAAAGCGGCATTTCCCTAATAAGCCCATTTTCCCGGGCGTTTTCATCGTAGAGGC 219
Db 14125 GTGTTTAAAGCGGCATTTCCCTAATAAGCCCATTTTCCCGGGCGTTTTCATCGTAGAGGC 14066

Qy 220 ATGGCGCAACGGGAGGTTTTCAGC 245
Db 14065 ATGGCGCAACGGGAGGTTTTCAGC 14040
```

ORIGIN

Query Match 12.7%; Score 61; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 4.8e-24;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 169 GGGCATTTCCTAATAAGCCCATTTCCCGGGCGTTTTCATCGTAGAGGGCATGGCGCAA 228
Db 169 GGGCATTTCCTAATAAGCCCATTTCCCGGGCGTTTTCATCGTAGAGGGCATGGCGCAA 228
Qy 229 A 229
Db 229 A 229

Search completed: January 12, 2006, 04:27:48
Job time : 3014 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 02:20:57 ; Search time 462 Seconds
(without alignments)
6924.356 Million cell updates/sec

Title: US-10-662-126-36
Perfect score: 480
Sequence: 1 atggaacaaagccatcaaaaa.....tgattgcagagagagattaa 480

Scoring table: OLIGO_NUC
Gapop 60:0 , Gapext 60:0

Searched: 4996997 seqs, 3332346308 residues

Word size : 24

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N_Geneseq_21.*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*
14: Geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	480	100.0	480	8	ACA34928
2	294	61.3	853	6	ABX66131 Helicobac
3	100	20.8	659	6	ABX66187 Helicobac
4	86	17.9	480	4	AAS53867 Helicobac

ALIGNMENTS

RESULT 1
ID ACA34928
XX ACA34928 standard; DNA; 480 BP.
XX ACA34928;
XX ACA34928;
DT 19-JUN-2003 (first entry)
XX Prokaryotic essential gene #16585.
DE
XX

Antisense; ds; prokaryotic essential gene; cell proliferation; drug design; gene.
XX Helicobacter pylori.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW; Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI: 2003-029926/02.
XX P-PSDB; ABU31058.
XX New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.
XX Claim 14; SEQ ID NO 22798; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 480 BP; 150 A; 93 C; 107 G; 130 T; 0 U; 0 Other;
XX Query Match 100.0%; Score 480; DB 8; Length 480;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-247;
XX Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAACAAAGCCATCAAACTTCGAATCTCAATTTTATAGAGCATATCTTACAAATT 60
DB 1 ATGGAACAAAGCCATCAAACTTCGAATCTCAATTTTATAGAGCATATCTTACAAATT 60

QY 61 CTACCTCACCCTATCCCATGCTTTTAGTGATAGAAATATAGATTACAAGCCAATAAA 120
DB |||||||
QY 121 AAAATGTGCGTTATAAGAAATACACTTTTAATAGAGCGTGTAAACGGGCATTTCCCT 180
DB |||||||
QY 181 AATAAGCCCATTTTCCCGGCGCTTTTGATCGTAGAGGCGATGCGCAACCGGAGGGTTT 240
DB |||||||
QY 241 TTAGCCTTCACTAGCTTGTGGGGTTTGACCTGAAATCCGCAAAACAAAATTCGTGTAT 300
DB |||||||
QY 301 TTCATGACGATTGATAAGTTAAATTCGCAATCCCTGTAAACCCAGGCGACAGATTAGAA 360
DB |||||||
QY 361 TACCATTAGAAAGTCTTAAAGCATAGGCGCATGATCTGGCAAGTGGTGGCAGGCTCAA 420
DB |||||||
QY 421 GTGGATGGCAAAAGTGTGCTGCAAGCCGAATTGAAAGCCATGATTGCAGAGAGATTAA 480
DB |||||||

RESULT 2

ABX66131
ID ABX66131 standard; DNA; 853 BP.

AC ABX66131;

DT 07-MAY-2003 (first entry)

DE Helicobacter pylori selected interacting domain (SID) DNA #730.

DE Protein-protein interaction; ulcer; selected interacting domain; SID;
KW gene; ds.

OS Helicobacter pylori.

PN WO200266501-A2.

PD 29-AUG-2002.

PF 28-DEC-2001; 2001WO-EP015428.

PR 02-JAN-2001; 2001US-0259302P.

XX (HYBR-) HYBRIGENICS.

PA (INSP) INST PASTEUR.

PI Legrain P, Rain J, Colland F, De Reuse H, Labigne A;

XX WPI; 2002-674910/72.

DR P-PSDB; ABUS1387.

PT New complexes of protein-protein interactions in Helicobacter pylori,
PT useful for identifying modulating compounds for treating or preventing
PT ulcers in mammals.

XX Claim 7; Page 263; 642pp; English.

XX The invention describes a complex of protein-protein interactions in
CC Helicobacter pylori selected from 421 complexes given in the
CC specification. The complex of protein-protein interactions are useful for
CC screening for agents which modulate the interaction of proteins.
CC Modulating compounds which binds to a targeted bacterial protein may be
CC used for treating or preventing ulcers in a human or animal. This
CC sequence encodes a selected interacting domain (SID), identified via

CC protein-protein interactions

SQ Sequence 853 BP; 258 A; 163 C; 167 G; 265 T; 0 U; 0 Other;

Query Match 61.3%; Score 294; DB 6; Length 853;
Best Local Similarity 100.0%; Pred. No. 1.9e-147; Indels 0; Gaps 0;
Matches 294; Conservative 0; Mismatches 0;

QY 1 ATGGAAACAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAAT 60
DB |||||||

QY 61 CTACCTCACCCTATCCCATGCTTTTAGTGGATAGAAATATAGATTACAAGCCAATAAA 120
DB |||||||

QY 62 CTACCTCACCCTATCCCATGCTTTTAGTGGATAGAAATATAGATTACAAGCCAATAAA 180
DB |||||||

QY 121 AAAATGTGCGTTATAAGAAATATCACTTTTAAATGAAGACGTGTAAACGGGCATTTCCCT 240
DB |||||||

QY 181 AATAAGCCCATTTTCCCGGCGCTTTTGATCGTAGAGGCGATGCGCAACCGGAGGGTTT 240
DB |||||||

QY 241 TTAGCCTTCACTAGCTTGTGGGGTTTGACCTGAAATCCGCAAAACAAAATC 294
DB |||||||

RESULT 3

ABX66187
ID ABX66187 standard; DNA; 659 BP.

AC ABX66187;

DT 07-MAY-2003 (first entry)

DE Helicobacter pylori selected interacting domain (SID) DNA #786.

DE Protein-protein interaction; ulcer; selected interacting domain; SID;
KW gene; ds.

OS Helicobacter pylori.

PN WO200266501-A2.

PD 29-AUG-2002.

PF 28-DEC-2001; 2001WO-EP015428.

PR 02-JAN-2001; 2001US-0259302P.

XX (HYBR-) HYBRIGENICS.

PA (INSP) INST PASTEUR.

PI Legrain P, Rain J, Colland F, De Reuse H, Labigne A;

XX WPI; 2002-674910/72.

DR P-PSDB; ABUS1443.

PT New complexes of protein-protein interactions in Helicobacter pylori,
PT useful for identifying modulating compounds for treating or preventing
PT ulcers in mammals.

XX Claim 7; Page 277; 642pp; English.

XX The invention describes a complex of protein-protein interactions in
CC Helicobacter pylori selected from 421 complexes given in the
CC specification. The complex of protein-protein interactions are useful for
CC screening for agents which modulate the interaction of proteins.
CC Modulating compounds which binds to a targeted bacterial protein may be
CC used for treating or preventing ulcers in a human or animal. This
CC sequence encodes a selected interacting domain (SID), identified via

CC protein-protein interactions

XX Sequence 659 BP; 200 A; 125 C; 123 G; 211 T; 0 U; 0 Other;

SQ Sequence 480 BP; 151 A; 94 C; 107 G; 128 T; 0 U; 0 Other;

Query Match 20.8%; Score 100; DB 6; Length 659;

Best Local Similarity 100.0%; Pred. No. 4.7e-43; Indels 0; Gaps 0;

Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAACAAAGCCATCAAACTTGGCAATCTCAATTTTATAGACATATCTTACAAATT 60

Db 560 ATGGAAACAAAGCCATCAAACTTGGCAATCTCAATTTTATAGACATATCTTACAAATT 619

QY 61 CTACCTCAGCGTATCCCATCTCTTTAGTGAAGAATTA 100

Db 620 CTACCTCAGCGTATCCCATCTCTTTAGTGAAGAATTA 659

RESULT 4

AAS53867

ID AAS53867 standard; DNA; 480 BP.

AC AAS53867;

DT 13-FEB-2002 (first entry)

XX Helicobacter pylori DNA for cellular proliferation protein #321.

DE Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

KW antibacterial; drug design.

XX Helicobacter pylori.

OS Helicobacter pylori.

XX WO200170955-A2.

PN 27-SEP-2001.

PD 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.

PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.

PR 22-DEC-2000; 2000US-0257931P.

PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX WPI: 2001-611495/70.

DR P-PSDB; AAU36008.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids.

XX Claim 27; SEQ ID NO 7504; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the genes,

CC their use in the discovery of novel antibiotics, the essential genes

CC themselves and the encoded proteins. The prokaryotes used are Escherichia

CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,

CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also

CC useful for the identification of potential new targets for antibiotic

CC development. The antisense nucleic acids can also be used to identify

CC proteins used in proliferation, to express these proteins, and to obtain

CC antibodies capable of binding to the expressed proteins. The proteins can

CC be used to screen compounds in rational drug discovery programmes. The

CC antisense nucleic acid sequence is also useful to screen for homologous

CC nucleic acids which are required for cell proliferation in a wide variety

CC of organisms. The present sequence encodes an essential prokaryotic

CC cellular proliferation protein. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 480 BP; 151 A; 94 C; 107 G; 128 T; 0 U; 0 Other;

SQ Sequence 480 BP; 151 A; 94 C; 107 G; 128 T; 0 U; 0 Other;

Query Match 17.9%; Score 86; DB 4; Length 480;

Best Local Similarity 100.0%; Pred. No. 1.6e-35;

Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 GTGTTTAAACGGGCATTTCCCTTAATAAGCCCATTTTCCGGGCGTTTTCATCGTAGAGGC 219

Db 160 GTGTTTAAACGGGCATTTCCCTTAATAAGCCCATTTTCCGGGCGTTTTCATCGTAGAGGC 219

QY 220 ATGGCGCAAAACGGGAGGGTTTTTAGC 245

Db 220 ATGGCGCAAAACGGGAGGGTTTTTAGC 245

Search completed: January 12, 2006, 03:37:28

Job time : 462 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 03:25:53 ; Search time 3072 Seconds
(without alignments)
7310.482 Million cell updates/sec

Title: US-10-662-126-36
Perfect score: 480
Sequence: 1 atggacaagaagccatcaaaa.....tgattgcagagagagattaa 480

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size: 24

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

- Database : EST:*
- 1: gb_est1:*
 - 2: gb_est2:*
 - 3: gb_est3:*
 - 4: gb_hic:*
 - 5: gb_est4:*
 - 6: gb_est5:*
 - 7: gb_est6:*
 - 8: gb_est7:*
 - 9: gb_gss1:*
 - 10: gb_gss2:*
 - 11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
-----	-----	-----	-----	-----	-----

No matches found

Search completed: January 12, 2006, 05:21:44
Job time : 3072 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 03:14:52 ; Search time 153 Seconds
(without alignments)
5576.664 Million cell updates/sec

Title: US-10-662-126-36
Perfect score: 480
Sequence: 1 atggaacaagccatcaaaaa.....tgattgcagagagagattaa 480

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1303057 seqs, 888780828 residues

Word size : 24

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

No matches found

Search completed: January 12, 2006, 04:30:27
Job time : 153 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 03:37:34 ; Search time 658 Seconds

(without alignments)
6032.373 Million cell updates/sec

Title: US-10-662-126-36

Perfect score: 480

Sequence: 1 atggaacaaagccatcaaaaa.....tgattgcagagagagattaa 480

Scoring table: Oligo NUC
Gapop 60.0 , Gapext 60.0

Searched: 9793542 seqs, 4134689005 residues

Word size : 24

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications NA.Main.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	480	100.0	480	7	US-10-282-122A-22798	Sequence 22798, A
2	480	100.0	480	9	US-10-662-126-36	Sequence 36, Appl
3	86	17.9	480	3	US-09-815-242-7504	Sequence 7504, Ap
4	86	17.9	480	7	US-10-335-977-2213	Sequence 2213, Ap
5	76	15.8	180	7	US-10-335-977-2212	Sequence 2212, Ap

ALIGNMENTS

RESULT 1
US-10-282-122A-22798
; Sequence 22798, Application US/10282122A
; Publication No. US20040029129A1

GENERAL INFORMATION:

- APPLICANT: Wang, Liangsu
- APPLICANT: Zamudio, Carlos
- APPLICANT: Malone, Cheryl
- APPLICANT: Haselbeck, Robert
- APPLICANT: Ohlsen, Kari
- APPLICANT: Zyskind, Judith
- APPLICANT: Wall, Daniel
- APPLICANT: Trawick, John

APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patentin version 3.1
SEQ ID NO 22798
LENGTH: 480
TYPE: DNA
ORGANISM: Helicobacter pylori
US-10-282-122A-22798

Query Match 100.0%; Score 480; DB 7; Length 480;
Best Local Similarity 100.0%; Pred. No. 1e-252;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT	60
DB	1	ATGGAACAAAGCCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT	60
QY	61	CTACCTCACCGCTATCCATGCTTTTAGTGCATAGAAATTATAGAGTTACAAGCCAATAA	120
DB	61	CTACCTCACCGCTATCCATGCTTTTAGTGCATAGAAATTATAGAGTTACAAGCCAATAA	120
QY	121	AAATTTGTCGCTTATAAGAAATATACATTTTAAAGAACGCTGTTTAAACGGGCATTTCCCT	180
DB	121	AAATTTGTCGCTTATAAGAAATATACATTTTAAAGAACGCTGTTTAAACGGGCATTTCCCT	180
QY	181	ATAAAGCCATTTTCCCGGCGTTTGATCGTAGAGGCGATGGCGCAACCGGAGGGTTT	240
DB	181	ATAAAGCCATTTTCCCGGCGTTTGATCGTAGAGGCGATGGCGCAACCGGAGGGTTT	240
QY	241	TTAGCCTTCACTAGCTTGTGGGGGTTTGACCTGAAATCGCCAAACAAAATCGTGAT	300
DB	241	TTAGCCTTCACTAGCTTGTGGGGGTTTGACCTGAAATCGCCAAACAAAATCGTGAT	300
QY	301	TTATGACGATTTGATAAGGTTAAATTCGCGATCCCTGTAACCCCGGCGCAGAGTTAGAA	360
DB	301	TTATGACGATTTGATAAGGTTAAATTCGCGATCCCTGTAACCCCGGCGCAGAGTTAGAA	360
QY	361	TACCAATTTAGAAGTCTTAAAGCATAAAGGGCATGATCTGGCAAGTGGTGGCACCGCTCAA	420
DB	361	TACCAATTTAGAAGTCTTAAAGCATAAAGGGCATGATCTGGCAAGTGGTGGCACCGCTCAA	420
QY	421	GTGGATGGCAAGTGTGCTGAAGCGCGAATTGAAGCCCATGATTGCAGAGAGATTAA	480
DB	421	GTGGATGGCAAGTGTGCTGAAGCGCGAATTGAAGCCCATGATTGCAGAGAGATTAA	480


```
RESULT 2
US-10-662-126-36
; Publication 36, Application US/10662126
; Publication No. US20050063987A1
; GENERAL INFORMATION:
; APPLICANT: Knapp, Bernhard
; APPLICANT: Hundt, Erika
; APPLICANT: Schmidt, Karl-Heinz
; TITLE OF INVENTION: Proteins, In Particular Membrane Proteins, of
; FILE OF INVENTION: Helicobacter Pylori, Their Preparation and Use
; FILE REFERENCE: CHIR-0340
; CURRENT APPLICATION NUMBER: US/10/662,126
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: 09/230,158
; PRIOR FILING DATE: 1999-01-19
; PRIOR APPLICATION NUMBER: PCT/IB97/00981
; PRIOR FILING DATE: 1997-07-25
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(480)
; OTHER INFORMATION: bp mature protein; 17 kD protein from Helicobacter
; OTHER INFORMATION: pylori
US-10-662-126-36

Query Match      100.0%; Score 480; DB 9; Length 480;
Best Local Similarity 100.0%; Pred. No. 1e-252;
Matches 480; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAACAAGGCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
DB 1 ATGGAACAAGGCATCAAACTTGCATCTCAATTTTATAGAGCATATCTTACAAATT 60
QY 61 CTACCTCAGCGTATCCCATGCTTTTAGTGATAGATATATAGATTACAGCCATATAA 120
DB 61 CTACCTCAGCGTATCCCATGCTTTTAGTGATAGATATATAGATTACAGCCATATAA 120
QY 121 AAAATGTGCTTTATAGAAATATCACTTTTAATGAAGACGTGTTTAACGGGCATTTCCCT 180
DB 121 AAAATGTGCTTTATAGAAATATCACTTTTAATGAAGACGTGTTTAACGGGCATTTCCCT 180
QY 181 AATAAGCCCATTTTCCCGGCGGTTTTCATCGTAGAGGCGATCGGCACAAACGGAGGTTT 240
DB 181 AATAAGCCCATTTTCCCGGCGGTTTTCATCGTAGAGGCGATCGGCACAAACGGAGGTTT 240
QY 241 TTAGCCTTCACCTAGCTTGTGGGGTTTGACCCCTGAATCGCCAAACAAAATCGTGAT 300
DB 241 TTAGCCTTCACCTAGCTTGTGGGGTTTGACCCCTGAATCGCCAAACAAAATCGTGAT 300
QY 301 TTATCATGACGATGATAGAGTTAAATTCGGCATCCCTGTAAACCCAGCGCAGATTAGAA 360
DB 301 TTATCATGACGATGATAGAGTTAAATTCGGCATCCCTGTAAACCCAGCGCAGATTAGAA 360
QY 361 TACCATTTAGAAGTCTTAAAGCATAAGGGCATGATCTGGCAAGTGGGTGGCAGCGTCAA 420
DB 361 TACCATTTAGAAGTCTTAAAGCATAAGGGCATGATCTGGCAAGTGGGTGGCAGCGTCAA 420
QY 421 GTGGATGGCAAGTGGTGGTGAAGCCGGAATTGAAGCCATGATTGACAGAGAGATTAA 480
DB 421 GTGGATGGCAAGTGGTGGTGAAGCCGGAATTGAAGCCATGATTGACAGAGAGATTAA 480

RESULT 3
US-09-815-242-7504
; Sequence 7504, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7504
; LENGTH: 480
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(480)
US-09-815-242-7504

Query Match      17.9%; Score 86; DB 3; Length 480;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 GTGTTTAAACGGGCATTTCCCTAATAAGCCCATTTTCCGGGGCGTTTTCATCGTAGAGGC 219
DB 160 GTGTTTAAACGGGCATTTCCCTAATAAGCCCATTTTCCGGGGCGTTTTCATCGTAGAGGC 219
QY 220 ATGCGCAAAACGGGAGGTTTTCAGC 245
DB 220 ATGCGCAAAACGGGAGGTTTTCAGC 245

RESULT 4
US-10-335-977-2213
; Sequence 2213, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 2213:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...480
; SEQUENCE DESCRIPTION: SEQ ID NO: 2213:
US-10-335-977-2213

Query Match      17.9%; Score 86; DB 7; Length 480;
Best Local Similarity 100.0%; Pred. No. 3.4e-36;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 GTGTTTAACGGGCATTTCCTAATAAGCCCATTTCCCGGGCGTTTTCGATCGTAGAGGC 219
Db 160 GTGTTTAACGGGCATTTCCTAATAAGCCCATTTCCCGGGCGTTTTCGATCGTAGAGGC 219

QY 220 ATGGCGCAACGGGAGGTTTTCAGC 245
Db 220 ATGGCGCAACGGGAGGTTTTCAGC 245

RESULT 5
US-10-335-977-2212
; Sequence 2212, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
```

```
;
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 2212:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...180
; SEQUENCE DESCRIPTION: SEQ ID NO: 2212:
US-10-335-977-2212

Query Match      15.8%; Score 76; DB 7; Length 180;
Best Local Similarity 100.0%; Pred. No. 1e-30;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 GGCATTTCCCTAATAAGCCCATTTTCGCGGGCGTTTTCGATCGTAGAGGCATGGCGCAA 229
Db 11 GGCATTTCCCTAATAAGCCCATTTTCGCGGGCGTTTTCGATCGTAGAGGCATGGCGCAA 70

QY 230 CGGAGAGGTTTTCAGC 245
Db 71 CGGAGAGGTTTTCAGC 86

Search completed: January 12, 2006, 05:43:16
Job time : 659 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2006, 03:29:48 ; Search time 619 Seconds
(without alignments)
627.605 Million cell updates/sec

Title: US-10-662-126-36
Perfect score: 480
Sequence: 1 atggaacaagccatcaaaaa.....tgattgcagagagagattaa 480

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 6038814 seqs, 404674181 residues

Word size : 2
Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications NA_New.*
1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2.*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3.*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

No matches found

Search completed: January 12, 2006, 05:32:10
Job time : 619 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:19:13 ; Search time 134 Seconds
(without alignments)
521.352 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 830

Sequence: 1 MQSHQNLOQSFIEHILQI.....QVDGKVVAEALKAMTAERD 159

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*
- 9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	830	100.0	159	6	ABU31058	Protein e
2	822	99.0	159	4	AU36008	Helicobac
3	420	50.6	146	6	ABU26345	Protein e
4	388	46.7	153	9	ABW91923	Abm91923 M. xanthu
5	386	46.5	145	8	ADN27276	Bacterial
6	377	45.4	140	8	ADS42226	Bacterial
7	376	45.3	138	8	ADS30389	Bacterial
8	352.5	42.5	152	8	ADS22800	Bacterial
9	349.5	42.1	162	7	ADF6342	Bacterial
10	348	41.9	172	8	ADS42775	Bacterial
11	347	41.8	139	6	ABU30001	Protein e
12	346	41.7	150	8	ADN17354	Bacterial
13	344.5	41.5	157	8	ADN17554	Bacterial
14	344.5	41.5	172	6	ABU40649	Protein e
15	342	41.2	147	8	ADS28964	Bacterial
16	341.5	41.1	160	7	ABO72472	Pseudomon
17	339.5	40.9	155	6	ABU22686	Protein e
18	339	40.8	138	8	ADS25172	Bacterial
19	339	40.8	138	8	ADS26072	Bacterial
20	339	40.8	138	8	ADS22670	Bacterial
21	339	40.8	138	8	ADS25610	Bacterial
22	338	40.7	146	6	ABU41863	Protein e
23	337	40.6	146	4	AU36344	Pseudomon
24	337	40.6	146	6	ABU38592	Protein e

25	337	40.6	146	7	ADE49809	Ade49809 (3R)-hydr
26	337	40.6	146	7	ADE49811	Ade49811 (3R)-hydr
27	332	40.0	154	8	ADS24815	Bacterial
28	332	40.0	181	6	ABU50043	Protein e
29	331	39.9	159	8	ADS27899	Bacterial
30	330	39.8	153	6	ABU49492	Protein e
31	329	39.6	144	6	ABU24817	Protein e
32	328	39.5	146	6	ABU39637	Protein e
33	323	38.9	138	8	ADN25760	Bacterial
34	323	38.9	140	8	ADS28418	Bacterial
35	323	38.9	151	4	AAU34452	E. coli c
36	323	38.9	151	6	ABU15025	Protein e
37	323	38.9	151	8	ADN18054	Bacterial
38	322	38.8	141	4	AAU35330	Enterococ
39	322	38.8	141	6	ABU29459	Protein e
40	322	38.8	142	7	ADH86668	Enterococ
41	321	38.7	151	6	ABU31679	Protein e
42	321	38.7	154	6	ABU21632	Protein e
43	320	38.6	141	6	ABU23918	Protein e
44	319.5	38.5	152	6	ABU39542	Protein e
45	319.5	38.5	155	6	ABU19696	Protein e

ALIGNMENTS

RESULT 1
ID ABU31058 standard; protein; 159 AA.
XX
AC ABU31058;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #16585.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Helicobacter pylori.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
WFI; 2003-029926/02.
N-PSDB; ACA34928.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 58982; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 159 AA;
Query Match 100.0%; Score 830; DB 6; Length 159;
Best Local Similarity 100.0%; Pred. No. 4.7e-91; Mismatches 0; Indels 0; Gaps 0;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIIELOQANKKIYVKNITFNEDVFNHGHP 60
DB 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIIELOQANKKIYVKNITFNEDVFNHGHP 60
QY 61 NKPIFGVLIVEGMAQTGGFLAFTSLMGFDPEIAKTIKIVYFMTIDKVKFRIPVTPGDRLE 120
DB 61 NKPIFGVLIVEGMAQTGGFLAFTSLMGFDPEIAKTIKIVYFMTIDKVKFRIPVTPGDRLE 120
QY 121 YHLEVLKHKGMIWQVGTAQVDGKVAEAEKAMIAERD 159
DB 121 YHLEVLKHKGMIWQVGTAQVDGKVAEAEKAMIAERD 159
RESULT 2
AAU36008
ID AAU36008 standard; protein; 159 AA.
XX
AC AAU36008;
DT
DT 14-FEB-2002 (first entry)
XX
DE Helicobacter pylori cellular proliferation protein #321.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX
OS Helicobacter pylori.
XX
XX WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009180.
XX
XX 21-MAR-2000; 2000US-0191078P.
XX 23-MAY-2000; 2000US-0206848P.
XX 26-MAY-2000; 2000US-0207272P.
XX 23-OCT-2000; 2000US-0242578P.
XX 27-NOV-2000; 2000US-0253625P.
XX 22-DEC-2000; 2000US-0257931P.
XX 16-FEB-2001; 2001US-0269308P.
XX

PA (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
XX N-PSDB; AAS3867.
XX
XX New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 11601; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 159 AA;
Query Match 99.0%; Score 822; DB 4; Length 159;
Best Local Similarity 98.7%; Pred. No. 4.3e-90; Mismatches 2; Indels 0; Gaps 0;
Matches 157; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIIELOQANKKIYVKNITFNEDVFNHGHP 60
DB 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIIELOQANKKIYVKNITFNEDVFNHGHP 60
QY 61 NKPIFGVLIVEGMAQTGGFLAFTSLMGFDPEIAKTIKIVYFMTIDKVKFRIPVTPGDRLE 120
DB 61 NKPIFGVLIVEGMAQTGGFLAFTSLMGFDPEIAKTIKIVYFMTIDKVKFRIPVTPGDRLE 120
QY 121 YHLEVLKHKGMIWQVGTAQVDGKVAEAEKAMIAERD 159
DB 121 YHLEVLKHKGMIWQVGTAQVDGKVAEAEKAMIAERD 159
RESULT 3
ABU26345
ID ABU26345 standard; protein; 146 AA.
XX
AC ABU26345;
XX
XX 19-JUN-2003 (first entry)
XX
XX Protein encoded by Prokaryotic essential gene #11872.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Campylobacter jejuni.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX

CC	improved plant properties, e.g. improved cold, heat or drought tolerance;
CC	tolerance to herbicides, extreme osmotic conditions, pathogens or pests;
CC	increased resistance to plant disease, better growth rate by modification
CC	of the cell cycle pathway with plant growth regulators, increased rate of
CC	homologous recombination, modified seed oil or protein yield and/or
CC	content, improved yield by modification of carbohydrate, nitrogen or
CC	phosphorus use and/or uptake, by modification of photosynthesis or by
CC	providing improved plant growth and development under at least one stress
CC	condition, improved lignin production or improved galactomannan
CC	production. This sequence represents a bacterial polypeptide used in the
CC	scope of the invention. Note: The sequence data for this patent did not
CC	form part of the printed specification but was obtained in electronic
CC	format from USPTO at seqdata.uspto.gov/sequence.html.
XX	
SQ	Sequence 138 AA;
	Query Match 45.3%; Score 376; DB 8; Length 138;
	Best Local Similarity 51.4%; Pred. No. 1.4e-36;
	Matches 72; Conservative 20; Mismatches 44; Indels 4; Gaps 2;
QY	14 IEHILQLPHRYPMLLVDRRIIEIQANKKIVAYKNITFNEVDVFNHGHPFNKPFPGLIVEG 73
DB	3 IGEILNLLPHRYPFLVLDRAVEIIPQKLTAYNQVTINEPFFNGHPFGHPVMPGVLLLEA 62
QY	74 MAQTGGFLAFTSLMGDPDPIAKTKIYVFMTIDKVFKRIPVTPGDRLEYHLVKHKGMW 133
DB	63 LAQATAILAVKS--ENMDP---SRKLTYLMGVDGARFRKPVLPGRDLQLBLETVVRHKGAVM 118
QY	134 QVGCTAQVDGKVVAAELXA 153
DB	119 KTKGLATVDGARVAEGEFLA 138
RESULT 8	
ADS22800	
ID	ADS22800 standard; protein; 152 AA.
XX	
AC	ADS22800;
XX	
DT	02-DEC-2004 (first entry)
XX	
DE	Bacterial polypeptide #11833.
XX	
KW	Recombinant DNA construct; transformed plant; improved plant property;
KW	cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW	pathogen tolerance; pest tolerance; plant disease resistance;
KW	cell cycle pathway modification; plant growth regulator;
KW	homologous recombination; seed oil yield; protein yield; carbohydrate;
KW	nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX	bacterial polypeptide.
OS	Bacteria.
PN	US2003233675-A1.
PX	
PD	18-DEC-2003.
XX	
PF	20-FEB-2003; 2003US-00369493.
XX	
PR	21-FEB-2002; 2002US-0360039P.
XX	
PA	(CAOY/) CAO Y.
PA	(HINK/) HINKLE G J.
PA	(SLAT/) SLATER S C.
PA	(CHEN/) CHEN X.
PA	(GOLD/) GOLDMAN B S.
XX	
PI	Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX	
DR	WPI; 2004-061375/06.
XX	
PT	New recombinant DNA construct comprising a promoter positioned to provide
PT	for expression of a polynucleotide encoding a polypeptide from a

PT microbial source, useful for producing plants with improved properties.
PS Claim 1; SEQ ID NO 11833; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 152 AA;
Query Match 42.5%; Score 352.5; DB 8; Length 152;
Best Local Similarity 46.2%; Pred. No. 1.1e-33;
Matches 66; Conservative 27; Mismatches 45; Indels 5; Gaps 1;
QY 17 ILQILPHRYPMMLVDRIELQANKKIVAYKNITFNEDVFNHGHPNKPFPGLVIVGMAQ 76
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 15 LMKLLPHRYPFLLMDRIIDIDGDSAIGIKNVITNEPFGHFPPEQVMPGVLIVEMAAQ 74
QY 77 TCGFLAFTSLWGFDPRIAKTKIVYFMTIDKVKPRIVPTGDRLEHYLHVKHGMVQVVG 136
||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 75 TAGAICIRSLGA-----SPSLVYFTIDNAKPKFVWPGDQKIHVKIKKRGNLLKFA 129
QY 137 GTAQVDGKVAEAEKAMIAERD 159
Db 130 CEALVDGTAKAAEASAMWVGTD 152
RESULT 9
ADF06342
ID ADF06342 standard; protein; 162 AA.
XX
AC ADF06342;
XX
DT 12-FEB-2004 (first entry)
XX
DE Bacterial polypeptide #2455.
XX
KW Proteus mirabilis infection; bacterial infection; antibacterial;
KW immunostimulant.
XX
OS Proteus mirabilis.
XX
FN US6605709-B1.
XX
PD 12-AUG-2003.
XX
PF 05-APR-2000; 2000US-00543681.
XX
PR 09-APR-1999; 99US-0128706P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX

PI Breton GL;
XX
DR WPI; 2003-895291/82.
DR N-PSDB; ADF02170.
XX
PT New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.
XX
PS Disclosure; SEQ ID NO 6627; 870pp; English.
XX
CC The invention relates to new Proteus mirabilis polypeptides and
CC polynucleotides. The invention also relates to antibodies against the
CC polypeptides, methods for producing the polypeptides, a method of
CC generating vaccines for immunising an individual against P. mirabilis, a
CC method for evaluating a compound for the ability to bind a P. mirabilis
CC polypeptide and a method for screening test compounds for anti-bacterial
CC activity. The polypeptides and polynucleotides are useful as molecular
CC targets for diagnosing, preventing and treating pathological conditions
CC resulting from bacterial infection, as reagents for diagnosis of
CC bacterial diseases, as components of antibacterial vaccines, as targets
CC for antibacterial drugs or as bio-control agents for plants. This
CC sequence represents a Proteus mirabilis polypeptide of the invention.
XX
SQ Sequence 162 AA;
Query Match 42.1%; Score 349.5; DB 7; Length 162;
Best Local Similarity 46.2%; Pred. No. 2.7e-33;
Matches 73; Conservative 27; Mismatches 47; Indels 11; Gaps 4;
QY 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIELQANKKIVAYKNITFNEDVFNHGHP 60
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 13 MSDNH-TLQ---IEEILDLLPHRYPFLLVDRIELQANKKIVAYKNITFNEDVFNHGHP 67
QY 61 NKPIFGVLIVGMAQVGGFLAFTSLWGFDPRIAKTKIVYFMTIDKVKPRIVPTGDRLE 120
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 68 GKPIFGVLIVGMAQVGGFLAFTSLWGFDPRIAKTKIVYFMTIDKVKPRIVPTGDRLE 123
QY 121 YHLEVLKHKGMVQVGGFLAFTSLWGFDPRIAKTKIVYFMTIDKVKPRIVPTGDRLE 158
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 124 LEVEFIRKRRGVARFVGAKVAGVGEIACEAEM--MCARR 159
RESULT 10
ADS42775
ID ADS42775 standard; protein; 172 AA.
XX
AC ADS42775;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #21205.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
FN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

QM protein - protein search, using sw model

Run on: January 10, 2006, 13:26:59 ; Search time 40 Seconds
(without alignments)
382.462 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 830

Sequence: 1 MEQSHQNLQSQFFIEHILQI.....QVDGKVAEAEKAMIAERD 159

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR 80:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	830	100.0	159	2 H64691	(3R)-hydroxymyrist
2	822	99.0	159	2 H71826	(3R)-hydroxymyrist
3	420	50.6	146	2 H81445	(3R)-hydroxymyrist
4	346	41.7	150	2 A70305	(3R)-hydroxymyrist
5	345.5	41.6	173	2 AB3356	(3r)-hydroxymyrist
6	339	40.8	154	2 AG2746	hypothetical prote
7	339	40.8	163	2 F97527	fabz protein (U516
8	337	40.6	146	2 E83190	(3R)-hydroxymyrist
9	332	40.0	181	2 AF0129	(3R)-hydroxymyrist
10	332	40.0	201	2 S35968	(3R)-hydroxymyrist
11	331	39.9	159	2 C87486	hypothetical prote
12	330	39.8	153	2 C82101	(3R)-hydroxymyrist
13	324	39.0	151	2 F90651	hypothetical prote
14	323	38.9	140	2 G84116	hydroxymyristoyl-(
15	323	38.9	151	2 D64742	(3R)-hydroxymyrist
16	323	38.9	141	2 F85502	hydroxymyristoyl-(
17	320	38.6	141	2 C97338	(3R)-hydroxymyrist
18	318	38.3	137	2 B72335	(3R)-hydroxymyrist
19	317	38.2	151	2 AE0530	hypothetical prote
20	316.5	38.1	145	2 A97701	hypothetical prote
21	313	37.7	144	2 AD1390	hydroxymyristoyl-(
22	313	37.7	144	2 AF1365	hydroxymyristoyl-(
23	311.5	37.5	145	2 B71708	(3R)-hydroxymyrist
24	309.5	37.3	140	2 H97919	(3R)-hydroxymyrist
25	309.5	37.3	148	2 G64180	(3R)-hydroxymyrist
26	308.5	37.2	140	2 B95049	hypothetical prote
27	308.5	37.2	144	2 AE6722	hypothetical prote
28	305.5	36.8	166	2 S75045	(3R)-hydroxymyrist
29	298.5	36.0	151	2 E86695	hypothetical prote

30	298	35.9	132	2 D70065	(3R)-hydroxymyrist
31	297.5	35.8	171	2 AH2089	(3R)-hydroxymyrist
32	288	34.7	134	2 C37083	probable (3R)-hydr
33	284	34.2	219	2 T50020	(3R)-hydroxymyrist
34	281.5	33.9	145	2 C84610	probable beta-hydr
35	277	33.4	146	2 H90002	hypothetical prote
36	276	33.3	149	2 G82000	(3R)-hydroxymyrist
37	275	33.1	149	2 D81228	(3R)-hydroxymyrist
38	269	32.4	160	2 C82731	myristoyl-acyl car
39	267.5	32.2	153	2 H86571	(3R)-hydroxymyrist
40	267.5	32.2	153	2 D72051	(3R)-hydroxymyrist
41	266.5	32.1	153	2 C71502	probable (3R)-hydr
42	265.5	32.0	153	2 H81661	(3R)-hydroxymyrist
43	252	30.4	251	2 D75439	(3R)-hydroxymyrist
44	164	19.8	156	2 B83881	(3R)-hydroxymyrist
45	148	17.8	159	2 AC2773	hypothetical prote

ALIGNMENTS

RESULT 1

H64691

(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Helicobacter pylori
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C;Accession: H64691
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997

A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: H64691

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-159 <TOM>

A;Cross-references: UNIPROT:O25928; UNIPARC:UP1000012A4AD; GB:AE000637; GB:AE000511; NID:

C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 100.0%; Score 830; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 7e-73;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEQSHQNLQSQFFIEHILQILPHRYPMLLVDRIIELOANKKIVAYKNITFNEDEVNGHFP 60

Db 1 MEQSHQNLQSQFFIEHILQILPHRYPMLLVDRIIELOANKKIVAYKNITFNEDEVNGHFP 60

Qy 61 NKPIFPGVLIVEGMAQTGGFLAFTSLMGDPDEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120

Db 61 NKPIFPGVLIVEGMAQTGGFLAFTSLMGDPDEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120

Qy 121 YHLEVLKHGMIMQVGCTAQDVGKVAEAEKAMIAERD 159

Db 121 YHLEVLKHGMIMQVGCTAQDVGKVAEAEKAMIAERD 159

RESULT 2

H71826

(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Helicobacter pylori
C;Species: Helicobacter pylori

A;Variety: strain J99

C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004

C;Accession: H71826

R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999

A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A;Reference number: A71800; MUID:99120557; PMID:9923682

A;Accession: H71826

A;Status: preliminary

A:Molecule type: DNA
A:Residues: 1-159 <ARN>
A:Cross-references: UNIPROT:Q9ZJL6; UNIPARC:UPI0000124AC; GB:AE001551; GB:AE001439; NID
A:Experimental source: strain J99
C:Genetics:
A:Gene: fabZ
C:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
C:Keywords: carbon-oxygen lyase; hydro-lyase; hydrolase

Query Match 99.0%; Score 822; DB 2; Length 159;
Best Local Similarity 98.7%; Pred. No. 4.1e-72;
Matches 157; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEQSHONLQSQFFIEHILQILPHRYPMLLDRIIELOQANKKIYAYKNITFNEDVFNHGHP 60
DB 1 MEQSHONLQSQFFIEHILQILPHRYPMLLDRIIELOQANKKIYAYKNITFNEDVFNHGHP 60

QY 61 NKPIPPGVNLIVEGMAQTGGLAFTSLWGFDPETAKTKIVYFMTIDKVKRIPVTPGDRLE 120
DB 61 NKPIPPGVNLIVEGMAQTGGLAFTSLWGFDPETAKTKIVYFMTIDKVKRIPVTPGDRLE 120

QY 121 YHLEVLKHKGMIVQVGTAGQVDGKVAEAEKAMIAERD 159
DB 121 YHLEVLKHKGMIVQVGTAGQVDGKVAEAEKAMIAERE 159

RESULT 3
H81445
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) Cj0273 [imported]
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: H81445
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
R.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: H81445
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-146 <PAR>
A:Cross-references: UNIPROT:Q9PIM2; UNIPARC:UPI0000124A2; GB:AL139074; GB:AL111168; NID
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: fabZ; Cj0273
C:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 50.6%; Score 420; DB 2; Length 146;
Best Local Similarity 56.8%; Pred. No. 2.4e-33;
Matches 83; Conservative 24; Mismatches 35; Indels 4; Gaps 2;

QY 14 IEHILQILPHRYPMLLDRIIELOQANKKIYAYKNITFNEDVFNHGHPNKPDPGVLIVEG 73
DB 4 VMQIQILPHRYPFLLDVKITELKVEWLGYNISIDHVFHMGHPGHPDPGVLIVEG 63

QY 74 MAQTGGLAFTSLWG-FDPRIAKTKIVYFMTIDKVKRIPVTPGDRLEHYHLEVLKHKGM 132
DB 64 MAQTGGLVAFESMEDKVPD---KSKVYFTGIDGAKFRNVRPGDRLDYEMSVYKNGNM 120

QY 133 WQVGGTAQVDGKVAEAEKAMIAER 158
DB 121 WIFGQAQVDGNLVAEAEKAMIVDK 146

RESULT 4
A70305
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: A70305
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov

Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: A70305
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-150 <AOQF>
A:Cross-references: UNIPROT:O66468; UNIPARC:UPI0000056250; GB:AE000671; NID:g2982793; PII
A:Experimental source: strain VF5
C:Genetics:
A:Gene: fabZ
C:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 41.7%; Score 346; DB 2; Length 150;
Best Local Similarity 50.7%; Pred. No. 3.5e-26;
Matches 71; Conservative 26; Mismatches 39; Indels 4; Gaps 2;

QY 14 IEHILQILPHRYPMLLDRIIELOQANKKIYAYKNITFNEDVFNHGHPNKPDPGVLIVEG 73
DB 3 IQEIMELPHRYPILLVDKILEIBEGKRIIIGLKNVSNPEPVQGHPPGFPPLPFGVYILEA 62

QY 74 MAQTGGLAFTSLWGFDPETAKTKIVYFMTIDKVKRIPVTPGDRLEHYHLEVLKHKGM 133
DB 63 MAQVGGILMIKSL---NLEIGKYAVV-FAGIDDRFKKPVTPGDDQIILEVVISLKALS 118

QY 134 QVGGTAQVDGKVAEAEKLA 153
DB 119 KMKGVAKVDGEVVAQATLMA 138

RESULT 5
AB3356
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) [imported] - Bruce
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 17-May-2002
C:Accession: AB3356
R:DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, I
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AB3356
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-173 <KUR>
A:Cross-references: UNIPARC:UPI0000057E2A; GB:AE008917; PIDN:AAL52013.1; PID:g17982777; C
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0832
A:Map position: 1
C:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 41.6%; Score 345.5; DB 2; Length 173;
Best Local Similarity 44.7%; Pred. No. 4.7e-26;
Matches 71; Conservative 24; Mismatches 59; Indels 5; Gaps 1;

QY 1 MEQSHONLQSQFFIEHILQILPHRYPMLLDRIIELOQANKKIYAYKNITFNEDVFNHGHP 60
DB 17 MSDNQTKLEAADIQALLAVLPHRYPFLLDRIIDVIGDVSATGINKVNTINBPHTGHP 76

QY 61 NKPIPPGVNLIVEGMAQTGGLAFTSLWGFDPETAKTKIVYFMTIDKVKRIPVTPGDRLE 120
DB 77 ENPIMPGLVIVEAMAQTAGAISL-----LQRTGTPGVVYFMTIDSAKFRPVPVPGDRLL 131

QY 121 YHLEVLKHKGMIVQVGTAGQVDGKVAEAEKAMIAERD 159
DB 132 LYVKIKORANISKYECVAEVDGKVAEAEVAMISVAD 170

RESULT 6
AG2746

hypothetical protein fabZ [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AG2746
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AG2746
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-154 <KUR>
A:Cross-references: UNIPROT:Q8UFL4; UNIPARC:UPI000012A49B; GB:AE008688; PIDN:AAL42389.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: fabZ
A:Map position: circular chromosome
C:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
Query Match 40.8%; Score 339; DB 2; Length 154;
Best Local Similarity 44.9%; Pred. No. 1.7e-25;
Matches 66; Conservative 29; Mismatches 36; Indels 16; Gaps 2;
QY 14 IEHILQILPHRYPMPLVDRIELQANKKIVAYKNITFNEVFNHGFNPKPIPPGVLIVEG 73
DB 14 ILEVKKLLPHRYPFLIDKILIEDGSSAIGIKVTVNEPHFTGHPDRPIPGVLIVEA 73
QY 74 MAQTGGFLAFTSLMGDFDPIATK-----IVYPMIDKVKFRIPVTPGDRLEHYHLEVLKH 128
DB 74 MAQTAGAIC-----ARNQGGGHLVYFMTIDNARFRPVPVGGDRLETHVVKQRQ 122
QY 129 KGMIVQVGTAQVDGKVVAAELKAMI 155
DB 123 RGNVFKFHCFAKVEGALVAEADVGM 149
RESULT 7
F97527
fabZ protein (U51683) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C:Accession: F97527
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Leppas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens C58
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: F97527
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-163 <KUR>
A:Cross-references: UNIPARC:UPI000016438F; GB:AE007869; PIDN:AAK87175.1; PID:gl5156449;
C:Genetics:
A:Gene: AGR_C2558
A:Map position: circular chromosome
C:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
Query Match 40.8%; Score 339; DB 2; Length 163;
Best Local Similarity 44.9%; Pred. No. 1.8e-25;
Matches 66; Conservative 29; Mismatches 36; Indels 16; Gaps 2;
QY 14 IEHILQILPHRYPMPLVDRIELQANKKIVAYKNITFNEVFNHGFNPKPIPPGVLIVEG 73
DB 23 ILEVKKLLPHRYPFLIDKILIEDGSSAIGIKVTVNEPHFTGHPDRPIPGVLIVEA 82
QY 74 MAQTGGFLAFTSLMGDFDPIATK-----IVYPMIDKVKFRIPVTPGDRLEHYHLEVLKH 128
DB 83 MAQTAGAIC-----ARNQGGGHLVYFMTIDNARFRPVPVGGDRLETHVVKQRQ 131

QY 129 KGMIVQVGTAQVDGKVVAAELKAMI 155
DB 132 RGNVFKFHCFAKVEGALVAEADVGM 158
RESULT 8
E83190
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1) PA3645 [imported]
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: E83190
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Briadman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, S.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: E83190
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-146 <STO>
A:Cross-references: UNIPROT:Q9HXY7; UNIPARC:UPI000012A4B1; GB:AE004784; GB:AE004091; NID:
A:Experimental source: strain PA01
C:Genetics:
A:Gene: fabZ; PA3645
A:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
C:Keywords: carbon-oxygen lyase; hydro-lyase
Query Match 40.6%; Score 337; DB 2; Length 146;
Best Local Similarity 49.3%; Pred. No. 2.5e-25;
Matches 72; Conservative 19; Mismatches 49; Indels 6; Gaps 3;
QY 14 IEHILQILPHRYPMPLVDRIELQ-ANKKIVAYKNITFNEVFNHGFNPKPIPPGVLIVE 72
DB 4 INEIREVLPHRYPFLVDVRLVDIEGKRIRAYKNVSNIEPFNFGHFEHPHPIPGVLIE 63
QY 73 GMAQTGGFLAFTSLMGDFDPIATKIVYFMTIDKVKFRIPVTPGDRLEHYHLEVLKHGM 132
DB 64 AMAQAAGILGFKML---DVKPADGTLYYFVGSGLKRFQPLPGDLQHLAKFISVKRSI 120
QY 133 KQVGGTAQVDGKVVAAELKAMTAER 158
DB 121 WKFDCHATVDKPKVCSAEI--ICAE 144
RESULT 9
AF0129
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1) (EC 4.2.1.1) [imported]
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AF0129
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, I.; Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AF0129
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-181 <KUR>
A:Cross-references: UNIPROT:Q8ZH57; UNIPARC:UPI000012A4C3; GB:AL590842; PIDN:CAC89897.1;
C:Genetics:
A:Gene: fabZ
A:Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
C:Keywords: carbon-oxygen lyase; hydro-lyase
Query Match 40.0%; Score 332; DB 2; Length 181;
Best Local Similarity 46.4%; Pred. No. 9.9e-25;
Matches 64; Conservative 27; Mismatches 43; Indels 4; Gaps 1;
QY 14 IEHILQILPHRYPMPLVDRIELQANKKIVAYKNITFNEVFNHGFNPKPIPPGVLIVEG 73
DB 11 ILEVKKLLPHRYPFLIDKILIEDGSSAIGIKVTVNEPHFTGHPDRPIPGVLIVEA 82


```

A;Accession: B3171
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-139,'DR',142-149,'RRDT' <COL>
A;Cross-references: UNIPARC:UPI0000178050; GB:M19334; GB:M18265; GB:M18266
C;Genetics:
A;Gene: fabZ
A;Start codon: TTG
C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match      38.9%; Score 323; DB 2; Length 151;
Best Local Similarity 46.4%; Pred. No. 5.9e-24;
Matches 64; Conservative 26; Mismatches 44; Indels 4; Gaps 1;

QY 14 IEHILQLPHRYPMLLVDRIIELOANKKIVAYKNITFNEVDVFNHGHPNKPPIPPGVLLIVEG 73
Db 10 IEIIELELPHRPFLLVDRLVDFEGRFLRAVKNVSVNEPFFQGHPPGKPIPPGVLLILEA 69

QY 74 MAQTGGFLAFTSLWGFDPPEIAKTKIVYFMFTIDKVFRIEFTPCDRLEYLEHLVKHKGMIV 133
Db 70 MAQATGILLAFKSVGKLEP---GELYYPAGIDEARPKRPVVPDQMI MEVTPKTRRGLT 125

QY 134 QVGGTAAQVDGKVVAAEL 151
Db 126 RPKGVALVDGKVVCEATM 143

Search completed: January 10, 2006, 13:35:48
Job time : 40 secs

```

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:26:04 ; Search time 160 Seconds
(without alignments)
701.119 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 830

Sequence: 1 MEQSHNLSQFFIEHLI.....QVDGKVAREALKAMTAERD 159

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	830	100.0	159	1	FABZ_HELPHY
2	822	99.0	159	1	FABZ_HELPHY
3	813	98.0	159	2	O5G940_HELPHY
4	509.5	61.4	169	1	FABZ_HELHP
5	488	58.8	151	1	FABZ_WOLSU
6	420	50.6	146	1	FABZ_CAMJE
7	420	50.6	146	1	FABZ_CAMJR
8	418	50.4	142	2	O4HJF5_CAMLA
9	417	50.2	142	2	O4HGA6_CAMCO
10	403	48.6	146	2	O4HQA3_CAMUP
11	396.5	47.8	156	2	Q4NZC4_9DELUT
12	389	46.9	153	1	FABZ_BRAJA
13	384	46.3	151	1	FABZ_RHIME
14	379	45.7	151	1	FABZ_RHOPA
15	375	45.2	150	1	FABZ_GROSL
16	365	44.0	154	1	FABZ_DSQVH
17	361.5	43.6	155	1	FABZ_BARVH
18	358.5	43.2	155	1	FABZ_BARHE
19	357	43.0	142	1	FABZ_CLOTE
20	355.5	42.8	145	1	FABZ_AZOSE
21	352.5	42.5	152	1	FABZ_RUILO
22	351.5	42.3	157	1	FABZ_BRUSU
23	351	42.3	150	1	FABZ_PHOLL
24	349.5	42.1	169	1	FABZ_GLUOX
25	348	41.9	150	1	FABZ_PROPR
26	347.5	41.9	164	1	FABZ_RALSO
27	346	41.7	150	1	FABZ_AQUAE
28	345.5	41.6	157	1	FABZ_AQUAE
29	345.5	41.6	157	1	FABZ_BRUME
30	344	41.4	151	1	FABZ_ERWCT
31	341	41.1	146	2	Q4ZWR7_PSESSY

32	340	41.0	150	2	O5E3F0_VIBF1	Q5e3f0 vibrio fisci
33	339.5	40.9	146	1	FABZ_METCA	Q604u1 methylococc
34	339.5	40.9	155	1	FABZ_BURMA	Q62jd5 burkholderi
35	339.5	40.9	155	1	FABZ_BURPS	Q63t23 burkholderi
36	339	40.8	154	1	FABZ_AGR5	Q8uf14 agrobacteri
37	338	40.7	146	1	FABZ_PSESM	Q886n2 pseudomonas
38	337	40.6	146	1	FABZ_PSEAE	Q9hxy7 pseudomonas
39	336	40.5	144	1	FABZ_BACLD	Q85e26 bacillus li
40	336	40.5	150	1	FABZ_VIBPA	Q87me8 vibrio para
41	334.5	40.3	467	1	LPXZ_CHLTE	Q8kx0 c lpxc/fabz
42	334	40.2	150	1	FABZ_MANSO	Q85ve3 manheimia
43	333	40.1	157	1	FABZ_SILPO	Q5lsu4 silicibacte
44	332	40.0	150	1	FABZ_DESPS	Q6aj07 desulfotale
45	332	40.0	150	1	FABZ_VIBVU	Q8dbf0 vibrio vuln

ALIGNMENTS

RESULT 1

ID	FABZ_HELPHY	STANDARD;	PRT;	159 AA.
AC	O25928;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)			
DE	((3R)-hydroxymyristoyl ACP dehydratase)			
GN	Names:fab2; OrderedLocusNames=HP1376;			
OS	Helicobacter pylori (Campylobacter pylori).			
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;			
OC	Helicobacteraceae; Helicobacter.			
OX	NCBI_TaxID=210;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC	STRAIN=26695 / ATCC 700392; PubMed=9252185; DOI=10.1038/41483;			
RX	MEDLINE=97394467; PubMed=9252185; DOI=10.1038/41483;			
RA	Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,			
RA	Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S.R.,			
RA	Dougherty B.A., Nelson K.E., Quackenbush J., Zhou L., Kirkness E.F.,			
RA	Peterson S.N., Loftus B.J., Richardson D.L., Dodson R.J., Khalak H.G.,			
RA	Glodek A., McKenney K., Fitzgerald L.M., Lee N., Adams M.D.,			
RA	Hickey E.K., Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D.,			
RA	Kelley J.M., Cotton M.D., Weidman J.F., Fujii C., Bowman C.,			
RA	Smith H.O., Fraser C.M., Venter J.C.;			
RT	"The complete genome sequence of the gastric pathogen Helicobacter			
RT	pylori.";			
RL	Nature 388:539-547(1997).			
CC	-I- FUNCTION: Involved in saturated fatty acids biosynthesis.			
CC	-I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	-I- SIMILARITY: Belongs to the thioester dehydratase family.			
CC	-----			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use as long as its content is in no way modified and this statement is not			
CC	removed.			
CC	-----			
CC	EMBL; AS000637; AAD08419.1; -; Genomic_DNA.			
DR	PIR; H64691; H64691.			
DR	TIGR; HPI376; -.			
DR	HMAP; MF_00406; -; 1.			
DR	InterPro; IPR010084; FabZ.			
DR	TIGRFAMs; TIGR01750; fabz; 1.			
KW	Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.			
FT	ACT SITE 58 58 By similarity.			
SQ	SEQUENCE 159 AA; 18196 MW; 715B6FBED872AE32 CRC64;			

Query Match 100.0%; Score 830; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.6e-71;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIELQANKKIYVAKNITFNEDVFNHGHP 60
Db 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIELQANKKIYVAKNITFNEDVFNHGHP 60
Qy 61 NKPIFPGVLIVEGMAQTGGFLAFTSLWGDFDEIATKIVVFMITDKVKFRIPVTPGDRLE 120
Db 61 NKPIFPGVLIVEGMAQTGGFLAFTSLWGDFDEIATKIVVFMITDKVKFRIPVTPGDRLE 120
Qy 121 YHLEVLKHKGMIWQVGTAQVDGKVAEAEELKAMIAERD 159
Db 121 YHLEVLKHKGMIWQVGTAQVDGKVAEAEELKAMIAERD 159

RESULT 2
FABZ_HELPJ STANDARD; PRT; 159 AA.
AC Q9ZJL6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 18-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=JHPI290;
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]
RX MEDLINE=99120557; PubMed=9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE001551; AAD06864.1; -; Genomic_DNA.
CC PIR; H71826; H71826.
CC HAMAP; MF 00406; -; 1.
CC InterPro; IPR010084; FabZ.
CC TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT SITE 58 By similarity.
SQ SEQUENCE 159 AA; 18210 MW; C0F16FBD868AE39 CRC64;

Query Match 99.0%; Score 822; DB 1; Length 159;
Best Local Similarity 98.7%; Pred. No. 9.2e-71;
Matches 157; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIELQANKKIYVAKNITFNEDVFNHGHP 60
Db 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIELQANKKIYVAKNITFNEDVFNHGHP 60
Qy 61 NKPIFPGVLIVEGMAQTGGFLAFTSLWGDFDEIATKIVVFMITDKVKFRIPVTPGDRLE 120
Db 61 NKPIFPGVLIVEGMAQTGGFLAFTSLWGDFDEIATKIVVFMITDKVKFRIPVTPGDRLE 120
Qy 121 YHLEVLKHKGMIWQVGTAQVDGKVAEAEELKAMIAERD 159
Db 121 YHLEVLKHKGMIWQVGTAQVDGKVAEAEELKAMIAERE 159
```

```
RESULT 3
Q5G940_HELPJ PRELIMINARY; PRT; 159 AA.
AC Q5G940;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE ((3R)-hydroxymyristoyl-acyl carrier protein dehydratase.
GN Name=fabZ;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SS1;
RX PubMed=15967411;
RA Liu W., Luo C., Han C., Peng S., Yang Y., Yue J., Shen X., Jiang H.;
RT "A new beta-hydroxyacyl-acyl carrier protein dehydratase (FabZ) from
RT Helicobacter pylori: Molecular cloning, enzymatic characterization,
RT and structural modeling.";
RL Biochem. Biophys. Res. Commun. 333:1078-1086(2005).
DR EMBL; AY725427; AAW22049.1; -; Genomic_DNA.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0016836; Fatty acid biosynthesis; IEA.
DR GO; GO:0006633; Fatty acid biosynthesis; IEA.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
SQ SEQUENCE 159 AA; 18184 MW; B7F16AAA78315B74 CRC64;

Query Match 98.0%; Score 813; DB 2; Length 159;
Best Local Similarity 97.5%; Pred. No. 6.7e-70;
Matches 155; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIELQANKKIYVAKNITFNEDVFNHGHP 60
Db 1 MEQSHQNLQSQFFIEHILQILPHRYPMMLVDRIELQANKKIYVAKNITFNEDVFNHGHP 60
Qy 61 NKPIFPGVLIVEGMAQTGGFLAFTSLWGDFDEIATKIVVFMITDKVKFRIPVTPGDRLE 120
Db 61 NKPIFPGVLIVEGMAQTGGFLAFTSLWGDFDEIATKIVVFMITDKVKFRIPVTPGDRLE 120
Qy 121 YHLEVLKHKGMIWQVGTAQVDGKVAEAEELKAMIAERD 159
Db 121 YHLEVLKHKGMIWQVGTAQVDGKVAEAEELKAMIAERE 159

RESULT 4
FABZ_HELPJ STANDARD; PRT; 169 AA.
AC Q7U3I9;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-acyl carrier protein] dehydratase (EC 4.2.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=HH1181;
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 31449 / 3B1;
RX MEDLINE=22709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;
RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,
RA Bell M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
```

FT	ACT SITE	49	49	By similarity.
SQ	SEQUENCE	151 AA;	17084 MW;	6869BA673438B628 CRC64;
Query Match				
	Best Local Similarity	58.8%;	Score 488;	DB 1; Length 151;
	Matches	90;	Conservative	24; Mismatches 32; Indels 0; Gaps 0;
Qy	14 IEHILQILPHRYPMLLVDRILIEIQANKKIVAYKNITFNEVDVFNHGFNPKPIFFGVLLIVEG	73		
Db	5 VQKIEILPHRRFPFLVDRVTALTSSIESIAYKNITINEEVFGQHPPIKVPVPGVLVIEG	64		
Qy	74 MAQTGGFLATSTLWGFDPETAKTKIVYFTWIDKVKRIPVTPGDRLEHYHLEV.LKHGMIW	133		
Db	65 MAQAGGYLAFVSMFGBEASNDHEKIVYFMSIDRAKFRVPVTPGDKLVYRLNVLKHGSIW	124		
Qy	134 QVGGTAQVDGKVVAAEELKAMIAERD	159		
Db	125 ILGRAYVDDKLVAAEELKAWADKE	150		
RESULT 6				
FABZ	CAMJE			
ID	-FABZ CAMJE	STANDARD;	PRT;	146 AA.
AC	Q9PMZ;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)			
DE	((3R)-hydroxymyristoyl ACP dehydratase).			
GN	Names-fabz; OrderedLocusNames=CJ0273;			
OS	Campylobacter jejuni.			
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;			
OC	Campylobacteraceae; Campylobacter.			
OX	NCBI_TaxID=197;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC	STRAIN=NTCT 11168;			
RX	MEDLINE=20150912; PubMed=10688204; DOI=10.1038/35001086;			
RA	Parkhill J., Wren B.W., Mungall K.L., Ketley J.M., Churcher C.M.,			
RA	Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,			
RA	Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,			
RA	Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,			
RA	Whitehead S., Barrrell B.G.;			
RT	"The genome sequence of the food-borne pathogen Campylobacter jejuni			
RT	reveals hypervariable sequences.";			
RL	Nature 403:665-668(2000).			
CC	-1- FUNCTION: Involved in saturated fatty acids biosynthesis.			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	-1- SIMILARITY: Belongs to the thioester dehydratase family.			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use as long as its content is in no way modified and this statement is not			
CC	removed.			
CC				
CC				
DR	EMBL: AL139074; CAB72741.1; -; Genomic_DNA.			
DR	PIR: H81445; H81445.			
DR	HAWAP; MF_00406; -; 1.			
DR	InterPro: IPR010084; FabZ.			
DR	TIGRFams; TIGR01750; fabZ; 1.			
KW	Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.			
FT	ACT SITE	48		
FT	SEQUENCE	146 AA;	16440 MW;	EFPA2656B12563A5 CRC64;
Query Match				
	Best Local Similarity	50.6%;	Score 420;	DB 1; Length 146;
	Matches	83;	Conservative	24; Mismatches 35; Indels 4; Gaps 2;
Qy	14 IEHILQILPHRYPMLLVDRILIEIQANKKIVAYKNITFNEVDVFNHGFNPKPIFFGVLLIVEG	73		
Db	4 VMQIQEILPHRRFPFLVDRVTALTSSIESIAYKNITINEEVFGQHPPIKVPVPGVLVIEG	63		

```
QY 74 MAQTGGFLAFTSLWG-FDPEIAKTKIVYFMTIDKVKFRIPVTPGDRLEVHLVLEKHKGM 132
DB 64 MAQTGGVLAFTSLWG-FDPEIAKTKIVYFMTIDKVKFRIPVTPGDRLEVHLVLEKHKGM 120
QY 133 WQVGGTAQVDGKVAEAEKAMIAER 158
DB 121 WIFKGAQFVDGKVAEAEKAMIVDK 146

RESULT 7
FABZ_CAMJR STANDARD; PRT; 146 AA.
ID QSHWJ3;
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
DE (3R)-hydroxymyristoyl-ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=CJ0322;
OS Campylobacter jejuni (strain RM1221).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=195099;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15660156; DOI=10.1371/journal.pbio.0030015;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Ravel J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural differences and novel potential virulence mechanisms
from the genomes of multiple Campylobacter species.";
RL PLOS Biol. 3:72-85(2005).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
DR EMBL; CP000025; AAW34912.1; -; Genomic_DNA.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 48 48 By similarity.
SQ SEQUENCE 146 AA; 16440 MW; EFA2656B12563A5 CRC64;

Query Match 50.6%; Score 420; DB 1; Length 146;
Best Local Similarity 56.8%; Pred. No. 3.3e-32;
Matches 83; Conservative 24; Mismatches 35; Indels 4; Gaps 2;

QY 14 IEHILPHRYPMLVDRIIELOANKKIVAYKNITFNEVDVFNHGHPNKPDPFGVLIVEG 73
DB 4 VMQIQILPHRYPFLVDKTELKKEVVLGYKNISIDSHVFNHGHPNKPDPFGVLIVEG 63

QY 74 MAQTGGFLAFTSLWG-FDPEIAKTKIVYFMTIDKVKFRIPVTPGDRLEVHLVLEKHKGM 132
DB 64 MAQTGGVLAFTSLWG-FDPEIAKTKIVYFMTIDKVKFRIPVTPGDRLEVHLVLEKHKGM 120
QY 133 WQVGGTAQVDGKVAEAEKAMIAER 158
DB 121 WIFKGAQFVDGKVAEAEKAMIVDK 146

RESULT 8
Q4HJP5_CAMLA PRELIMINARY; PRT; 142 AA.
ID Q4HJP5_CAMLA PRELIMINARY; PRT; 142 AA.
AC Q4HGA6;
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Beta-hydroxyacyl-(Acyl-carrier-protein) dehydratase FabZ
DE (EC 4.2.1.-).
GN Name=fabZ; ORFNames=CC00342;
OS Campylobacter coli RM2228.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306254;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM2228;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
genomes of multiple Campylobacter species.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
```

```
Q4HJP5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Beta-hydroxyacyl-(Acyl-carrier-protein) dehydratase FabZ
DE (EC 4.2.1.-)
GN Name=fabZ; ORFNames=CLA0287;
OS Campylobacter lari RM2100.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306263;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM2100;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
genomes of multiple Campylobacter species.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL; AAPK01000004; EAL54516.1; -; Genomic_DNA.
DR KW Lyase.
SQ SEQUENCE 142 AA; 16066 MW; D4D1DBA5660BC286 CRC64;

Query Match 50.4%; Score 418; DB 2; Length 142;
Best Local Similarity 56.6%; Pred. No. 5e-32;
Matches 81; Conservative 27; Mismatches 31; Indels 4; Gaps 2;

QY 17 ILQILPHRYPMLVDRIIELOANKKIVAYKNITFNEVDVFNHGHPNKPDPFGVLIVEGMAQ 76
DB 3 IQXILPHRYPFLVDKTELKKEVVLGYKNISIDSHVFNHGHPNKPDPFGVLIVEGMAQ 62

QY 77 TGGFLAFTSLWGDPPEI-AKTKIVYFMTIDKVKFRIPVTPGDRLEVHLVLEKHKGMIVQV 135
DB 63 TGGVLAFTSLWGDPPEI-AKTKIVYFMTIDKVKFRIPVTPGDRLEVHLVLEKHKGMIVQV 119

QY 136 GGTAAQVDGKVAEAEKAMIAER 158
DB 120 EGKAFVDGKVAEAEKAMIVDK 142

RESULT 9
Q4HGA6_CAMCO PRELIMINARY; PRT; 142 AA.
ID Q4HGA6_CAMCO PRELIMINARY; PRT; 142 AA.
AC Q4HGA6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Beta-hydroxyacyl-(Acyl-carrier-protein) dehydratase FabZ
DE (EC 4.2.1.-).
GN Name=fabZ; ORFNames=CC00342;
OS Campylobacter coli RM2228.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306254;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM2228;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
genomes of multiple Campylobacter species.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
```

```
CC preliminary data.
DR EMBL; AAFJ01000003; EAL56875.1; -; Genomic_DNA.
KW Lyase.
SQ SEQUENCE 142 AA; 15937 MW; 2E000F62AC59B479 CRC64;

Query Match 50.2%; Score 417; DB 2; Length 142;
Best Local Similarity 57.3%; Pred. No. 6.2e-32;
Matches 82; Conservative 25; Mismatches 32; Indels 4; Gaps 2;

QY 17 ILQILPHRYPMMLVDRIELQANKKIYAKNITFNEVDVFNHGFNPKPIPPGVLIIVEGMAQ 76
DB 3 IQEILPHRYPFLLDVKITELKVGVEVRGKKNISIDSHVFMGHFGHPPIYPGVLLLEGMAQ 62

QY 77 TGGFLAFTSLWGFDPRI-AKTIVVYFMTIDKVKFRIPVTPGDRLEVHLVLEKHKGMIVQV 135
DB 63 TGGVLAFESM---DAKVDPKSVVYFTGIDGAKFRNPVREGDLRYEMQVVKRGNWIF 119

QY 136 GGTAAVDGKVAEAEELKAMIAER 158
DB 120 KGQAFVDGMLVAEAEELKAMIVDK 142

RESULT 10
Q4HQ3 CAMUP PRELIMINARY; PRT; 146 AA.
AC Q4HQ3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Beta-hydroxyacyl-(Acyl-carrier-protein) dehydratase FabZ
DE (EC 4.2.1.1.-).
GN Name=fabZ; ORFNames=CUP1147;
OS Campylobacter upsaliensis RM3195.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=308264;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM3195;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Raeko D.A.,
RA Jacques R.J., Burkinac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R.R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
RT genomes of multiple Campylobacter species."
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFJ01000007; EAL52872.1; -; Genomic_DNA.
KW Lyase.
SQ SEQUENCE 146 AA; 16591 MW; 8F1CD2BA02891959 CRC64;

Query Match 48.6%; Score 403; DB 2; Length 146;
Best Local Similarity 55.5%; Pred. No. 1.4e-30;
Matches 81; Conservative 24; Mismatches 37; Indels 4; Gaps 2;

QY 14 IEHILQILPHRYPMMLVDRIELQANKKIYAKNITFNEVDVFNHGFNPKPIPPGVLIIVEG 73
DB 4 IMQIQEILPHRYPFLLDVKITELKVGVEVRGKKNISIDSHVFMGHFGHPPIYPGVLLLEG 63

QY 74 MAQTGGFLAFTSLWGFDPRI-AKTIVVYFMTIDKVKFRIPVTPGDRLEVHLVLEKHKGM 132
DB 64 MAQTGGVLAFESM---DNKVNPKSVVYFTGIDGAKFRNPVREGDLRYEMQVVKRGNAL 120

QY 133 WOVGTAAVDGKVAEAEELKAMIAER 158
DB 121 WIFESKAFVEENLVAEAEELKAMIVDK 146

RESULT 11
QANZC4_9DEL
```

```
ID QANZC4_9DEL PRELIMINARY; PRT; 156 AA.
AC QANZC4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Beta-hydroxyacyl-(Acyl-carrier-protein) dehydratase FabZ.
GN ORFNames=AdehDRAFT_3875;
OS Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cystobacterineae; Myxococcaceae; Anaeromyxobacter.
OX NCBI_TaxID=290397;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHD01000002; EAL80756.1; -; Genomic_DNA.
SQ SEQUENCE 156 AA; 17343 MW; 38620F1678976741 CRC64;

Query Match 47.8%; Score 396.5; DB 2; Length 156;
Best Local Similarity 53.1%; Pred. No. 6.4e-30;
Matches 76; Conservative 24; Mismatches 40; Indels 3; Gaps 1;

QY 17 ILQILPHRYPMMLVDRIELQANKKIYAKNITFNEVDVFNHGFNPKPIPPGVLIIVEGMAQ 76
DB 16 IQEILPHRYPFLLDVVRVEFEAHRRLVAKGVTVNNEPFFQGHFPAQVPWPGVLLLEALAQ 75

QY 77 TGGFLAFTSLWGFDPRI-AKTIVVYFMTIDKVKFRIPVTPGDRLEVHLVLEKHKGMIVQV 136
DB 76 AAALLATMSL---KPDEVKDKITVLMGIDGARFRRPVVPGDRLEVEVTRKQKGVWQKT 132

QY 137 GTAQVDGKVAEAEELKAMIAERD 159
DB 133 GVARVDGQVVAEAEFMAMLADRE 155

RESULT 12
FABZ_BRAJA STANDARD; - PRT; 153 AA.
ID FABZ_BRAJA
AC Q89KQ3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-(acyl carrier protein) dehydratase (EC 4.2.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydrase).
GN Name=fabZ; OrderedLocusNames=bll14851;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
```


RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RL Bradyrhizobium japonicum USDA110.";
CC DNA Res. 9:189-197(2002).

CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity)

CC -!- SIMILARITY: Belongs to the thioester dehydratase family.

CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC -----
CC

DR HMAP; BA000040; BAC50116.1; -; Genomic_DNA.

DR HAMAP; MF 00406; -; 1.

DR InterPro; IPR010084; FabZ.

DR TIGRFAMs; TIGR01750; fabz; 1.

KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.

FT ACT SITE 58 By similarity.

SQ SEQUENCE 153 AA; 17178 MW; 337FB1B3D5C1FFDC CRC64;

Query Match 46.9%; Score 389; DB 1; Length 153;

Best Local Similarity 49.3%; Pred. No. 3.3e-29;

Matches 71; Conservative 30; Mismatches 39; Indels 4; Gaps 1;

QY 14 IEHILQLPHRYPMMLVDRIIELOANKKIVAYKNITFNEDVFNHGFNPKPIPPGVLIVEG 73

DB 14 INAILQLPHRYPMMLIDRVINRADYSGIKNVTNFEPAFQGHPPRPVYGVMMIEA 73

QY 74 MAQTGFLAFTSLMGDPDPIAKTKIVYFTIDKVPRIPTGDRLEYLEVLEKGM1W 133

DB 74 MAQTAGVIGKISVEGTE-----KPRVYFLTIDCKFKPVLPGDTIEYMRSLGRKTMW 129

QY 134 QVGGTAQVDGKVVAAELKAMIAE 157

DB 130 WPHGDAKNGVQVAAEDVAGMLTD 153

RESULT 13

FABZ RHIME STANDARD; PRT; 154 AA.

AC Q92Q46;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)

DE ((3R)-hydroxymyristoyl ACP dehydratase).

GN Name=fabZ; OrderedLocusNames=R01504; ORFNames=SMC02092;

OS Rhizobium melliloti (Sinorhizobium melliloti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Sinorhizobium.

OX NCBI_TaxID=382;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=1021;

RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;

RA Capela D., Barloy-Hubler P., Gouzy J., Bothe G., Ampe F., Batut J.,

RA Boistard P., Becker A., Boudry M., Cadieu E., Dreano S., Gloux S.,

RA Godrie T., Goffeau A., Kahn D., Kles E., Lelaure V., Masuy D.,

RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsberger U.,

RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,

RT "Analysis of the chromosome sequence of the legume symbiont

RT Sinorhizobium melliloti strain 1021";

RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).

CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- SIMILARITY: Belongs to the thioester dehydratase family.

CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC

DR HMAP; AL591787; CAC46083.1; -; Genomic_DNA.

DR HAMAP; MF 00406; -; 1.

DR InterPro; IPR010084; FabZ.

DR TIGRFAMs; TIGR01750; fabz; 1.

KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.

FT ACT SITE 57 By similarity.

SQ SEQUENCE 154 AA; 16994 MW; 6E1BB0EA6F8CDF40 CRC64;

Query Match 46.3%; Score 384; DB 1; Length 154;

Best Local Similarity 50.0%; Pred. No. 1e-28;

Matches 75; Conservative 24; Mismatches 37; Indels 14; Gaps 2;

QY 14 IEHILQLPHRYPMMLVDRIIELOANKKIVAYKNITFNEDVFNHGFNPKPIPPGVLIVEG 73

DB 13 IQBILRLPHRYPFLVLDRIIEIDDDNSAIGIKNVTANEPHFTGHFPEKPIPPGVLLIEG 72

QY 74 MAQTGFLAFTSLMGDPDPEAKT-----KIVYFTIDKVKRIPVTPGDRLEYLEVLEKHK 129

DB 73 MAQTAGAIC-----ARKTGIGSNLYVFTIDNARFRKPVVPGDRVEFHVVKQQR 122

QY 130 GMTWQVGGTAQVDGKVVAAELKAMIAERD 159

DB 123 GNIWKHCDKADKVDGQIVAEADIGAMIVSKE 152

RESULT 14

FABZ RHOPA STANDARD; PRT; 151 AA.

AC P61454;

DT 05-JUL-2004 (Rel. 44, Created)

DT 05-JUL-2004 (Rel. 44, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)

DE ((3R)-hydroxymyristoyl ACP dehydratase).

GN Name=fabZ; OrderedLocusNames=RPA2912;

OS Rhodopseudomonas palustris.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Bradyrhizobiaceae; Rhodopseudomonas.

OX NCBI_TaxID=1076;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=CGA009 / ATCC BAA-98;

RX PubMed=14704707; DOI=10.1038/nbt923;

RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,

RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,

RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,

RA Harrison F.H., Gibson J., Harwood C.S.;

RT "Complete genome sequence of the metabolically versatile

RT photosynthetic bacterium Rhodopseudomonas palustris.";

RL Nat. Biotechnol. 22:55-61(2004).

CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- SIMILARITY: Belongs to the thioester dehydratase family.

CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC -----
CC

DR HMAP; BX572602; CAE28353.1; -; Genomic_DNA.

DR HAMAP; MF 00406; -; 1.

DR InterPro; IPR010084; FabZ.

DR TIGRFAMs; TIGR01750; fabz; 1.

KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.

FT ACT SITE 56 By similarity.

SQ SEQUENCE 151 AA; 17017 MW; 8F9989D2D9B9BFC8 CRC64;

Query Match 45.7%; Score 379; DB 1; Length 151;

Best Local Similarity 48.6%; Pred. No. 3e-28;

Matches 70; Conservative 29; Mismatches 41; Indels 4; Gaps 1;

Search completed: January 10, 2006, 13:34:58
Job time : 161 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:29:09 ; Search time 46 Seconds
(without alignments)
285.770 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 830

Sequence: 1 MQSHQNLQSFIEHILQI.....QVHGKVAEELKAMIAERD 159

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Issued Patents AA:*
- 1: /cgn2_6/prodata/1/iaa/5_COMB.pep.*
 - 2: /cgn2_6/prodata/1/iaa/6_COMB.pep.*
 - 3: /cgn2_6/prodata/1/iaa/H_COMB.pep.*
 - 4: /cgn2_6/prodata/1/iaa/PCRTUS_COMB.pep.*
 - 5: /cgn2_6/prodata/1/iaa/RE_COMB.pep.*
 - 6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	388	46.7	153	2	US-09-902-540-11122
2	349.5	42.1	162	2	US-09-543-681A-6627
3	341.5	41.1	160	2	US-09-252-991A-21218
4	322	38.8	142	2	US-09-134-000C-4553
5	317	38.2	186	2	US-09-489-039A-12393
6	311.5	37.5	140	2	US-09-583-110-3906
7	311.5	37.5	146	2	US-09-107-433-2892
8	308.5	37.2	140	2	US-09-196-388-2
9	308.5	37.2	160	2	US-10-089-019-18
10	294	35.4	185	2	US-09-540-236-3473
11	283	34.1	156	2	US-09-134-001C-4234
12	278	33.5	166	2	US-10-089-019-10
13	277	33.4	146	2	US-09-339-614-2
14	271	32.7	163	2	US-09-328-352-7242
15	267.5	32.2	153	2	US-09-198-452A-689
16	267.5	32.2	153	2	US-09-438-185A-653
17	233	28.1	117	2	US-09-107-532A-5188
18	193.5	23.3	132	2	US-09-902-540-11447
19	152	18.3	196	2	US-09-543-681A-6762
20	116.5	14.0	149	2	US-09-902-540-11443
21	98	11.8	173	2	US-09-543-681A-7869
22	98	11.8	174	2	US-09-252-991A-25897
23	85.5	10.3	180	2	US-09-248-796A-19898
24	80.5	9.7	489	2	US-09-134-001C-4902
25	79	9.5	467	2	US-09-248-796A-16476
26	77	9.3	316	2	US-09-248-796A-18830
27	75.5	9.1	121	2	US-09-602-777A-322

28	75.5	9.1	121	2	US-09-602-777A-324	Sequence 324, App
29	75	9.0	260	2	US-09-270-767-33502	Sequence 33502, A
30	75	9.0	260	2	US-09-270-767-48719	Sequence 48719, A
31	73.5	8.9	1002	2	US-10-290-579A-187	Sequence 187, App
32	73.5	8.9	1003	2	US-10-290-579A-193	Sequence 193, App
33	73	8.8	195	2	US-09-489-039A-12028	Sequence 12028, A
34	72.5	8.7	311	2	US-09-489-039A-9786	Sequence 9786, Ap
35	72.5	8.7	1002	2	US-10-290-579A-184	Sequence 184, App
36	72.5	8.7	1003	2	US-10-290-579A-188	Sequence 188, App
37	72	8.7	422	2	US-09-248-796A-16762	Sequence 16762, A
38	71.5	8.6	247	2	US-09-173-300-45	Sequence 27781, A
39	71.5	8.6	424	2	US-09-173-300-45	Sequence 45, Appl
40	71.5	8.6	424	2	US-10-027-450-45	Sequence 45, Appl
41	71.5	8.6	466	2	US-10-104-047-3687	Sequence 3687, Ap
42	71.5	8.6	850	2	US-09-952-060-4	Sequence 4, Appl
43	71.5	8.6	875	2	US-09-952-060-8	Sequence 8, Appl
44	71.5	8.6	1003	1	US-07-743-357-10	Sequence 10, Appl
45	71.5	8.6	1005	2	US-10-290-579A-191	Sequence 191, App

ALIGNMENTS

RESULT 1

US-09-902-540-11122
; Sequence 11122, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 11122
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; US-09-902-540-11122

Query Match	46.7%	Score 388:	DB 2:	Length 153;
Best Local Similarity	50.0%	Pred. No. 8.5e-38;		
Matches	73;	Conservative	23;	Mismatches 46;
			Indels	4;
			Gaps	2;
Qy	14	IEHILQLPHRYPMMLVDRIIELOANKKIVAYKNITFEDVFNHGFNPKPIPPGVLIVRG	73	
Db	3	IGETLNLPHRYPLLDVDRVEIIPGQKLTAYKNVTINEPFNFGHFGHPGVPVPGVLILEA	62	
Qy	74	MAQTGGPLATSLMGDPFEIAKTVYFMTIDKVKFRIPVTPGDRLEVHLVKHKGMW	133	
Db	63	LAQATAILAKYS-ENMDP---SRKLTVMGVDGARFRKPLPGDRLQLEIEVVRHKGAW	118	
Qy	134	QVGGTAQVDGKVVAEELKAMIAERD	159	
Db	119	KTKGLATVDGARVAEGEFLATVVVDK	144	

RESULT 2

US-09-543-681A-6627
; Sequence 6627, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709-1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05

```

; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6627
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6627

```

Query Match	42.1%	Score	349.5	DB 2	Length	162
Best Local Similarity	46.2%	Pred. No.	3.2e-33			
Matches	73	Conservative	27	Mismatches	47	Indels
					11	Gaps
QY	1	MEOSHQNLSQSFETHIQLPHRYPMLLVDRIITELQANKKIYAKNTTFNEDVFNGHFP	60			
Db	13	MSDNH-TLQ---IEELDLPHRYPELLVDRIIDFEEGKFLRAVKNSVFNPTFQGHFP	67			
QY	61	NKEPIFGVLIVEGAQTGGFLAFTSLMGFDPEIAKTIVYFMTIDVKFRFIPVTPGDLR	120			
Db	68	GKPIFGVLILEMAQATGILAFKSVGKLEP---GELYYPAAIDGARFKFRLPGDQMV	123			
QY	121	YHLEVLKHGMIVQVGTAQVDGKVAEAEELKAMIAER	158			
Db	124	LEYEFIKERGVARFKGAVKVDGSIACEAEN--MCARR	159			

```

RESULT 3
US-09-252-991A-21218
; Sequence 21218, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21218
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21218

```

Query Match	41.1%;	Score	341.5;	DB 2;	Length	160;			
Best Local Similarity	47.5%;	Pred. No.	2.8e-32;						
Matches	75;	Conservative	21;	Mismatches	53;	Indels	9;	Gaps	4;
Qy	2	EQSHQNLOSQFFTEHLILQILPHRYPMLLVDRIIEIQ-ANKKIVAYKNITFNEDVFNGHFP	60						
Db	9	EQAPPNNMD---INEIREYLPHTYPTLLVQKRVLEDIEGKKIRAYKNVSNINEPFFNGHFP	65						
Qy	61	NKPIFPGLVIVEGMAQQGGFLAFTSLMGFDPDEIAKTIVYEMTITDKVKFRIPVPTPGDRLE	120						
Db	66	EHPITMGVLLIEAWQAAGILGFQWL---DVKPADGTILYFVSGDKLFRQPVLPQDQLQ	122						
Qy	121	YHLEVLKHGMIMWGVTGAQVDGKVKVAEAEELKAMIAER	158						
Db	123	LHAKFISVKSIWKFCDCHATVDDKPVCSAET--ICAER	158						

RESULT 4
US-09-134-000C-4553
; Sequence 453, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AN
; TITLE OF INVENTION: ENTEROCOCCUS FAEC

```

; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4553
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4553

```

Query Match	38.8%;	Score 322;	DB 2;	Length 142;
Best Local Similarity	50.0%;	Pred. NO. 4.7e-30;		
Matches	71;	Conservative 19;	Mismatches 46;	Indels 6; Gaps 3;
QY	14	IEHILQLPHRYPMLLAVDRIIELOANKKIVAYKNITFNEDVFGNHPNKPIFPQVLIVEG	73	
Db	6	IYEIQLPHRYPFLDLDSVEEVIPEGVRVAKKVNTVNEQVFQGHFPGNPVLFGVLIIES	65	
QY	74	MAQTGGFLAFTSLMGDPDIATKIIVFYMTIDIKVKFERIPVTGPDRLLEYHLLEVHKGMW	133	
Db	66	LAQAQA-VALLSW----PEF-KGKTAYFGGLDKAKFRQKVTTPGDTLILEVELLKVRSAG	119	
QY	134	QVGCGTAQVDGKVVAAELXAMI	155	
Db	120	MKGKGVAKVNGKKVAEALTFMI	141	

RESULT 5

```

US-09-489-039A-12393
; Sequence 12393. Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AL
; TITLE OF INVENTION: PNEUMONIAE FOR DIA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,0
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,74
; PRIOR FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12393
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12393

```

Query Match	38.2%	Score 317;	DB 2;	Length 186;
Best Local Similarity	45.7%;	Pred. No. 2.7e-29;		
Matches	63;	Conservative 26;	Mismatches 45;	Indels 4; Gaps 1;
Qy	14	IEHILQILPHRYDMLAVDRITIELQANKKIWAYKNITFNEDVFNGHPNKKPIFPQGVLI	VBG	73
Db	45	IEKIVELLPHRYEFLVDRVDFEGRFURAVKNVSVNEEFFQGHFPCKILPGVLILEA	104	
Qy	74	MAQTGGFLAFTSLWGDPDEIAKTKIYVFMTIDKVKFRIPVTPGDRLEHYHLEVLXHKGMIV	133	
Db	105	MAQATGILAFSGVKLEP---GELYFFAGIDEARFKRPVVPQGMIMEVTFEKTRRGLT	160	
Qy	134	QVGGTAQVDGKVVAAEAL	151	
Db	161	RFKGVALVDGKVVCEATM	178	

RESULT 6
US-09-583-110-3906
; Sequence 3906, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: LYNN Doucette-Stamm et al.


```

/ APPLICANT: DEWOLF, WALTER E. JR
/ APPLICANT: KALLENDER, HOWARD
/ APPLICANT: LONSDALE, JOHN T.
/ TITLE OF INVENTION: METHODS FOR MAKING AND USING FATTY ACID
/ TITLE OF INVENTION: SYNTHESIS PATHWAY REAGENTS
/ FILE REFERENCE: GM50068
/ CURRENT APPLICATION NUMBER: US/10/089,019
/ CURRENT FILING DATE: 2002-03-25
/ PRIOR APPLICATION NUMBER: PCT/US00/29451
/ PRIOR FILING DATE: 2000-10-26
/ PRIOR APPLICATION NUMBER: 60/161,775
/ PRIOR FILING DATE: 1999-10-27
/ NUMBER OF SEQ ID NOS: 37
/ SOFTWARE: FastSeq for Windows version 4.0
/ SEQ ID NO 18
/ LENGTH: 160
/ TYPE: PRT
/ ORGANISM: Streptococcus pneumoniae
/ US-10-089-019-18

```

	Query Match	37.2%	Score 308.5;	DB 2;	Length 160;
	Best Local Similarity	47.8%;	Pred. No. 2.2e-28;		
	Matches 66;	Conservative 22;	Mismatches 43;	Indels 7;	Gaps 3;
Qy	14	IEHLOILPHRYPMLVDRILIELQANKKIIVAYKNITNEDEVENGHENKPIFGVGLIVEG	73		
Db	24	IOGKEALPHRYPMLLVDRVLEV-SETVIAIKVNTINEPFFNGHFFQYPMWPGVGLIMEA	82		
Qy	74	MAOTGGGFLATSLWGFDPPEIAKTIKVYFMTIDVKVFRIPVTPGDRLEVHLVLELKHGMIV	133		
Db	83	LAQTAGVLELSK-----PE-NKGKLVFVAGMDVKVFKKQVVPGLQVMTATFVKRRGTIA	136		
Qy	134	QVGGTAQVDGKVVAAEEL	151		
Db	137	VVEAKAEVDGKLAASGTL	154		

RESULT 10
 US-09-540-236-3473
 ; Sequence 3473, Application US/09540236
 ; Patent No. 6673910
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2005-001
 ; CURRENT APPLICATION NUMBER: US/09/540,236
 ; CURRENT FILING DATE: 2000-04-04
 ; NUMBER OF SEQ ID NOS: 3840
 ; SEQ ID NO 3473
 ; LENGTH: 185
 ; TYPE: prt
 ; ORGANISM: M.catarrhalis
 US-09-540-236-3473

Query Match	35.4%;	Score 294;	DB 2;	Length 185;
Best Local Similarity	45.6%;	Pred. No. 1.4e-26;		
Matches	67;	Conservative 23;	Mismatches 47;	Indels 10; Gaps 4;
Qy	16	HILQI---LPHRYPMLAVDRILIELQANKKIYAVKNITFNEDVFNGHPNKPPIPGVLVIVE	72	
Db	39	HYNQIKHYLPHRYPMLIDRVITACKPNEWITGYKNISINEELFNHGFDPNIMPGVLOVE	98	
Qy	73	GMAQTGPFAPFTSLWGDPDEIATKIVV-FMTIDKVFRIPIVTPGDRLEHYLHVKHEKM	131	
Db	99	ANAQLSGILGFISA---GQTADDDGYLYLPAGVDKVRPKVVTSGDQLVIRSKVLMNKRD	154	
Qy	132	IWQVGGTAQVDQKGVVAEELKAMIAER	158	
Db	155	IVKPECTAHVDGOLTCGAELI--MIARO	179	

RESULT 11

```

US-09-134-001C-4234
; Sequence 4234, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4234
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4234

```

Query Match	34.1%	Score 283;	DB 2;	Length 156;
Best Local Similarity	40.7%;	Pred. No. 2.2e-25;		
Matches	61;	Conservative 21;	Mismatches 62;	Indels 6;
Gaps	1;			
QY	8	LSQOFFEHILOILPHRYPMILVDRIITELQANKKIIVAYKNIITNEDVFNHGHPNKPFPFG	67	
DB	12	METTFYDNQIKIIPHROPFLIDKIIVEYBEGKRCVGLKQSGNEPFGQHGFYAVNWP	71	
QY	68	VLIVEGMAQTGGFLAFTSLWGFDPDEIARTKIIVYPMTIDKVKFRIPVTPFGDRLEHYHLEVK	127	
DB	72	VLITELAQTGAVALNS-----EENKGKIALFAGIDKCFKQKQVVGDTLMLEVEITK	125	
QY	128	HKGMIWQVGGTQVDGKVAEAEUKAMIAE	157	
DB	126	IKGPIGKGSAKATVDGOLACSELTFAIQD	155	

RESULT 12
US-10-089-019-10
; Sequence 10, Application US/10089019
; Patent No. 6951729
; GENERAL INFORMATION:
; APPLICANT: DEWOLF, WALTER E. JR
; APPLICANT: KALLENDER, HOWARD
; APPLICANT: LONSDALE, JOHN T.
; TITLE OF INVENTION: METHODS FOR MAKING AND USING FATTY ACID
; TITLE OF INVENTION: SYNTHESIS PATHWAY REAGENTS
; FILE REFERENCE: GW50068
; CURRENT APPLICATION NUMBER: US/10/089,019
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: PCT/US00/29451
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/161,775
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-089-019-10

	Query Match	33.5%	Score 278;	DB 2;	Length 166;	
	Best Local Similarity	39.1%;	Pred. No. 9.2e-25;			
	Matches	59;	Conservative	24;	Mismatches	62; Indels 6; Gaps 1
QY	7	NLQSFHIIQLILPHRYMLLVDRITIELQANKKI VAYKNITFNEVDVFNHGFPNKPFFP	66			
		:::::	:::::	:::::	:::::	:::::
Dd	20	HMETFDYNQIKIIPHROPFLIDKVVYESEGRCAVKQVSGNEPFQFGHPFYAVMP	79			
		:::::	:::::	:::::	:::::	:::::
QY	67	GVLIVEGMAQTGGFLAFTSLMGDPFEIAKTIVVFMTIDKVKFRIPVTTPGDRLEVHVHLV	126			
		:::::	:::::	:::::	:::::	:::::

Db 80 GVLITLALAQGTGAVALINS-----BENKGIKALFAGIDKCRFKRQVVGDTLTLEVEIT 133
Qy 127 KHGMIMWQVGTQAVDQGVKVAEALKAMIAE 157
Db 134 KIKGPIKGNKATVDGQLACSCELTFAIQD 164

RESULT 13

US-09-339-614-2
; Sequence 2, Application US/09339614
; Patent No. 6489139
; GENERAL INFORMATION:
; APPLICANT: Kallender, Howard
; APPLICANT: Van Horn, Stephanie
; APPLICANT: Warren, Patrick V.
; APPLICANT: Lonsdale, John
; TITLE OF INVENTION: fab2
; FILE REFERENCE: GM10228
; CURRENT APPLICATION NUMBER: US/09/339,614
; CURRENT FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-339-614-2

Query Match 33.4%; Score 277; DB 2; Length 146;
Best Local Similarity 39.3%; Pred. No. 1e-24;
Matches 59; Conservative 23; Mismatches 62; Indels 6; Gaps 1;
Qy 8 LOSQFFIEHILQILPHRYPMMLVDRIELQANKKIYAVKNITFNEDVFNHGHFPNKPFP 67
Db 1 METIPDYNQIKQIIPHQFPLLDKVVVEEGQRCVAIKQVSGNEPFFQGHFPPEYAVMPG 60
Qy 68 VLIIVGMAQTGFLAFTSLWGFDPDEIAKTIYVFMITDKVKFRIPVTPGDRLEHYHLEVLK 127
Db 61 VLIITLALAQGTGAVALINS-----BENKGIKALFAGIDKCRFKRQVVGDTLTLEVEITK 114
Qy 128 HKGMIWQVGTQAVDQGVKVAEALKAMIAE 157
Db 115 IKGPIKGNKATVDGQLACSCELTFAIQD 144

RESULT 14

US-09-328-352-7242
; Sequence 7242, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7242
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7242

Query Match 32.7%; Score 271; DB 2; Length 163;
Best Local Similarity 44.6%; Pred. No. 6e-24;
Matches 62; Conservative 19; Mismatches 52; Indels 6; Gaps 3;
Qy 14 IEHILQILPHRYPMMLVDRIELQANKKIYAVKNITFNEDVFNHGHFPNKPFPGLVIVEG 73
Db 20 IQTIRQYLPFRYPFLVDVRVTVDN-SIVGYKNVNSINEEFLQGHFPEYPIIMPGLVIVEA 78
Qy 74 MAQTGGFLAFTSLWGFDPDEIAKT-KIVYFMITDKVKFRIPVTPGDRLEHYHLEVLKHGMI 132

Db 79 LAQVSGVLGFI---MNNETPKGSLFLFAGARVRFKQVVGDLVLKSELVQMQRGI 134
Qy 133 WQVGTQAVDQGVKVAEAL 151
Db 135 YKYNCTASVDGIVAATAEI 153

RESULT 15

US-09-198-452A-689
; Sequence 689, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prever
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 689
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-689

Query Match 32.2%; Score 267.5; DB 2; Length 153;
Best Local Similarity 40.9%; Pred. No. 1.4e-23;
Matches 65; Conservative 25; Mismatches 54; Indels 15; Gaps 7;
Qy 7 NLOSQFFIEHILQILPHRYPMMLVDRII--ELQANKKIYAVKNITFNEDVFNHGHFPNKPDI 64
Db 2 NQPSVIKLRLLDILLPHRYPEFLVDKVLSDIEA-RSITAQKNVTINEPFFMGHFPNAPI 60
Qy 65 FPGVLIIVGMAQTGFLAFTSLWGFDPDEIAKTK-IVYFMITDKVKFRIPVTPGD--RLEY 121
Db 61 MPGLVILLEALAAQAGVLI-----GLVLENDNRKRIALFLGIQAKFQAVRPGDVLTLQA 115
Qy 122 HLEVLRKHK-GMIWQVGTQAVDQGVKVAEALKAMIAERD 159
Db 116 DFSLISSKSGKAW---AQARVDSQIVTEAELSPALVDKE 151

Search completed: January 10, 2006, 13:36:46
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:29:45 ; Search time 61 Seconds
(without alignments)
1089.096 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 830

Sequence: 1 MEQSHQNLSQFFIEHLIQL.....QVDGKVVAELKAMIAERD 159

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/prodata1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/prodata1/pubpaa/US10A_PUBCOMB.pep:*
- 5: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/prodata1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	830	100.0	159	US-10-282-122A-58982	Sequence 58982, A
2	830	100.0	159	US-10-662-126-37	Sequence 37, Appl
3	830	100.0	159	US-10-965-006-11	Sequence 11, Appl
4	822	99.0	159	US-09-815-242-11601	Sequence 11601, A
5	822	99.0	159	US-10-335-977-6975	Sequence 6975, Ap
6	420	50.6	146	US-10-282-122A-54269	Sequence 54269, A
7	386	46.5	145	US-10-369-493-9929	Sequence 9929, Ap
8	377	45.4	140	US-10-369-493-20656	Sequence 20656, A
9	376	45.3	138	US-10-369-493-19422	Sequence 19422, A
10	352.5	42.5	152	US-10-369-493-11833	Sequence 11833, A
11	348	41.9	172	US-10-369-493-21205	Sequence 21205, A
12	347	41.8	139	US-10-282-122A-57925	Sequence 57925, A
13	346	41.7	150	US-10-369-493-7	Sequence 7, Appl
14	344.5	41.5	157	US-10-369-493-207	Sequence 207, Appl
15	344.5	41.5	172	US-10-282-122A-68573	Sequence 68573, A
16	342	41.2	147	US-10-369-493-17997	Sequence 17997, A
17	339.5	40.9	155	US-10-282-122A-50610	Sequence 50610, A
18	339	40.8	138	US-10-369-493-11703	Sequence 11703, A
19	339	40.8	138	US-10-369-493-14205	Sequence 14205, A
20	339	40.8	138	US-10-369-493-14643	Sequence 14643, A
21	339	40.8	138	US-10-369-493-15105	Sequence 15105, A
22	338	40.7	146	US-10-282-122A-69787	Sequence 69787, A
23	337	40.6	146	US-09-815-242-11937	Sequence 11937, A
24	337	40.6	146	US-10-282-122A-66516	Sequence 66516, A
25	337	40.6	146	US-10-965-006-2	Sequence 2, Appl
26	337	40.6	146	US-10-965-006-4	Sequence 4, Appl
27	332	40.0	154	US-10-369-493-13848	Sequence 13848, A

ALIGNMENTS

RESULT 1

US-10-282-122A-58982

; Sequence 58982, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; PRIOR FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 58982

; LENGTH: 159

; TYPE: PRT

; ORGANISM: Helicobacter pylori

US-10-282-122A-58982


```
Query Match      100.0%; Score 830; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 9.5e-85;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEQSHQNLSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAKNITFNEDEVNGHPP 60
DB 1 MEQSHQNLSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAKNITFNEDEVNGHPP 60
QY 61 NKPIFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
DB 61 NKPIFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
QY 121 YHLEVLKHKGMIWQVGTAQVDGKVVAEAEKAMIAERD 159
DB 121 YHLEVLKHKGMIWQVGTAQVDGKVVAEAEKAMIAERD 159

RESULT 2
US-10-662-126-37
; Sequence 37, Application US/10662126
; Publication No. US20050063987A1
; GENERAL INFORMATION:
; APPLICANT: Knapp, Bernhard
; APPLICANT: Hundt, Erika
; TITLE OF INVENTION: Schmidt, Karl-Heinz
; TITLE OF INVENTION: Proteins, In Particular Membrane Proteins, of
; FILE REFERENCE: Helicobacter Pylori, Their Preparation and Use
; FILE REFERENCE: CHIR-0340
; CURRENT APPLICATION NUMBER: US/10/662,126
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: 09/230,158
; PRIOR FILING DATE: 1999-01-19
; PRIOR APPLICATION NUMBER: PCT/IB97/00981
; PRIOR FILING DATE: 1997-07-25
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-10-662-126-37

Query Match      100.0%; Score 830; DB 5; Length 159;
Best Local Similarity 100.0%; Pred. No. 9.5e-85;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEQSHQNLSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAKNITFNEDEVNGHPP 60
DB 1 MEQSHQNLSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAKNITFNEDEVNGHPP 60
QY 61 NKPIFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
DB 61 NKPIFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
QY 121 YHLEVLKHKGMIWQVGTAQVDGKVVAEAEKAMIAERD 159
DB 121 YHLEVLKHKGMIWQVGTAQVDGKVVAEAEKAMIAERD 159

RESULT 3
US-10-965-006-11
; Sequence 11, Application US/10965006
; Publication No. US2005021462A1
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, ALED
; APPLICANT: DHARMSI, AKIL
; APPLICANT: VEDADI, MASOUD
; APPLICANT: DOMAGALA, MEGAN
; APPLICANT: KIMBER, MATTHEW
; APPLICANT: VALLEE, FRANCOIS
; TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM PSEUDOMONAS AERUGINOSA
; FILE REFERENCE: IPT-257.01
; CURRENT APPLICATION NUMBER: US/10/965,006
```

```
; CURRENT FILING DATE: 2004-10-14
; PRIOR APPLICATION NUMBER: PCT/CA03/00560
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: 60/373,321
; PRIOR FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 11
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-10-965-006-11

Query Match      100.0%; Score 830; DB 5; Length 159;
Best Local Similarity 100.0%; Pred. No. 9.5e-85;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEQSHQNLSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAKNITFNEDEVNGHPP 60
DB 1 MEQSHQNLSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAKNITFNEDEVNGHPP 60
QY 61 NKPIFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
DB 61 NKPIFGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
QY 121 YHLEVLKHKGMIWQVGTAQVDGKVVAEAEKAMIAERD 159
DB 121 YHLEVLKHKGMIWQVGTAQVDGKVVAEAEKAMIAERD 159

RESULT 4
US-09-815-242-11601
; Sequence 11601, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Orlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11601
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11601

Query Match      99.0%; Score 822; DB 3; Length 159;
Best Local Similarity 98.7%; Pred. No. 7.5e-84;
Matches 157; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 MEQSHONLQSQFFIEHILQILPHRYPMLLVDRILIELQANKKIYVAKNITFNEVDVFNHGHP 60
Db 1 MEQSHONLQSQFFIEHILQILPHRYPMLLVDRILIELQANKKIYVAKNITFNEVDVFNHGHP 60
QY 61 NKPIPPGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
Db 61 NKPIPPGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
QY 121 YHLEVLKHGMIWQVGGTAQVDGKVAEAEELKAMIAERD 159
Db 121 YHLEVLKHGMIWQVGGTAQVDGKVAEAEELKAMIAERE 159

RESULT 5

US-10-335-977-6975
; Sequence 6975, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSER: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997

ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 6975:
SEQUENCE CHARACTERISTICS:
LENGTH: 159 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori

FEATURE:
NAME/KEY: misc.feature
LOCATION: (B) LOCATION 1...159
SEQUENCE DESCRIPTION: SEQ ID NO: 6975:
US-10-335-977-6975

Query Match 99.0%; Score 822; DB 4; Length 159;
Best Local Similarity 98.7%; Pred. No. 7.5e-84;
Matches 157; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEQSHONLQSQFFIEHILQILPHRYPMLLVDRILIELQANKKIYVAKNITFNEVDVFNHGHP 60
Db 1 MEQSHONLQSQFFIEHILQILPHRYPMLLVDRILIELQANKKIYVAKNITFNEVDVFNHGHP 60
QY 61 NKPIPPGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120

Db 61 NKPIPPGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
QY 121 YHLEVLKHGMIWQVGGTAQVDGKVAEAEELKAMIAERD 159
Db 121 YHLEVLKHGMIWQVGGTAQVDGKVAEAEELKAMIAERE 159

RESULT 6

US-10-282-122A-54269
; Sequence 54269, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54269
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
US-10-282-122A-54269

Query Match 50.6%; Score 420; DB 4; Length 146;
Best Local Similarity 56.8%; Pred. No. 9e-39;
Matches 83; Conservative 24; Mismatches 35; Indels 4; Gaps 2;
QY 14 IEHILQILPHRYPMLLVDRILIELQANKKIYVAKNITFNEVDVFNHGHPFNKPIPPGVLIVEG 73
Db 4 VMOIQEILPHRYPFLVLVDKTELKVEVLGYKNISISDHVFMGHPFGHPDIYGVLLIEG 63
QY 74 MAQTGGFLAFTSLWG-FDPEIAKTKIVYFMTIDKVKFRIPVTPGDRLEIYHLEVLKHGMI 132
Db 64 MAQTGGVLAFESMEDKVP---KSKVYFTGIDGAKFRNVPFGDRLDYEMSVVKNRGM 120
QY 133 WQVGGTAQVDGKVAEAEELKAMIAER 159
Db 121 WIFKGOAFVDGNLVAEAEELKAMIVDK 146


```

; ORGANISM: Mesorhizobium loti
US-10-369-493-11833

Query Match      42.5%; Score 352.5; DB 4; Length 152;
Best Local Similarity 46.2%; Pred. No. 3.6e-31;
Matches 66; Conservative 27; Mismatches 45; Indels 5; Gaps 1;

QY 17 ILQLPHRYPMMLVDRIELQANKKIYVKNITFNEVDVFNHGHPNKPFPFGLVIVEGMAQ 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 15 LMKLLPHRYPFLMIDRIIDIGDSDSAIGIKNVTINEPHFQGHPEQVMPFGLVIVEAMAQ 74
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 77 TGGFLAFTSLWGFDPETAKTKIVYFMTIDKVKPRIPVTPGDRLEVHLVLEKHKGMWQVG 136
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 75 TAGAICIRSLGA-----SKSLVFLTIDNAKFRKVPVPGDQLKIHKVKKIKRGNLKPFA 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 137 GTAQVDGKVAEAEAKAMIAERD 159
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 CEALVDGTGAABAEIASAMVMTGD 152
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
US-10-369-493-21205
; Sequence 21205, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21205
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
; FEATURE:
; LOCATION: (1)..(172)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-21205

Query Match      41.9%; Score 348; DB 4; Length 172;
Best Local Similarity 48.3%; Pred. No. 1.4e-30;
Matches 70; Conservative 25; Mismatches 44; Indels 6; Gaps 2;

QY 14 IEHILQILPHRYPMMLVDRIELQANKKIYVKNITFNEVDVFNHGHPNKPFPFGLVIVEG 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 31 IEHILQILPHRYPFLVDRIELQANKKIYVKNITFNEVDVFNHGHPNKPFPFGLVILEA 90
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 74 MAQTGGFLAFTSLWGFDPETAKTKIVYFMTIDKVKPRIPVTPGDRLEVHLVLEKHKGMW 133
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 91 MAQTGILAFKSV-----GSLAPGKLYPFAIDGARFKRPVLPDGMVLEVEFKERRGVA 146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 134 QVGGTAQVDGKVAEAEAKAMIAER 158
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 147 RFGVAKVDGEVACEAEM--MCARR 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-10-282-122A-57925
; Sequence 57925, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl

```

```

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57925
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Enterococcus faecium
US-10-282-122A-57925

Query Match      41.8%; Score 347; DB 4; Length 139;
Best Local Similarity 53.5%; Pred. No. 1.3e-30;
Matches 76; Conservative 16; Mismatches 44; Indels 6; Gaps 2;

QY 14 IEHILQILPHRYPMMLVDRIELQANKKIYVKNITFNEVDVFNHGHPNKPFPFGLVIVEG 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 IOBKIEIIPHYPMMLTDRVEEMVGERIVAKNVTINEPFFQGHPEEPVMPGLVILEA 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 74 MAQTGGFLAFTSLWGFDPETAKTKIVYFMTIDKVKPRIPVTPGDRLEVHLVLEKHKGMW 133
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 MAQAGA-VALLSLBQP-----KGTAYFGGLDKAKFRKKVTPGDTLYLEVLKVKASAG 116
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 134 QVGGTAQVDGKVAEAEAKAMIAER 155
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 117 IGKGIKVDGKVAEAEALTFMI 138
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-10-369-493-7
; Sequence 7, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493

```

```

; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 7
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Aquifex aeolicus
US-10-369-493-7

Query Match 41.7%; Score 346; DB 4; Length 150;
Best Local Similarity 50.7%; Pred. No. 1.9e-30;
Matches 71; Conservative 26; Mismatches 39; Indels 4; Gaps 2

QY 14 IEHILQLPHRYPMLLVDRIIEIQANKKIVAYKNTTFNEDVFNGHFPNKPFPFGLVIRG 73
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3 IQEIMEILPHRYPIILLVDKILEIEEGKRIIGLKNVSVNEPVFGHFPFGLFPGVYILEA 62
QY 74 MAQTGGGLAFTSLMGDFDEIAKTKIVYPMTTDKVKFRIPVPTPGDRLEHYHLEVLKHKGIW 133
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
63 MAQVGGGILMIKSL---NLEIGKYAVV-PAGIDDAARFKKPVYPGDLILELEVISLKKALS 118
QY 134 QVGGTAQVDGKVAEAEELKA 153
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
119 KMGVAKVGDGEVVAQATLMA 138

RESULT 14
US-10-369-493-207
; Sequence 207, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTIVITY
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 207
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(157)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-207

Query Match 41.5%; Score 344.5; DB 4; Length 157;
Best Local Similarity 46.2%; Pred. No. 3e-30;
Matches 73; Conservative 27; Mismatches 47; Indels 11; Gaps 1

QY 1 MEQSHQLQSOFIEHILQILPHRYPMLLVDRIIEIQANKKIVAYKNTTFNEDVFNGHFP 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
8 MSDNH-TLQ---IEEILKLLPHRYPFLLVDRLDFEKGKFLRAVKNTVNEPFFQGHFP 62
QY 61 NKPIFPFGLVIRGMAQTGGFLAFTSLMGDFDEIAKTKIVYPMTTDKVKFRIPVPTPGDRLE 120
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
63 GKPIFPFGLVIRLEAMAQATGILAFKTVGTLEP----GELYFFAIDGARPKRPVLPDQMI 118
QY 121 YHLEVLKHKGIWVGGTGAQVDGKVAEAEELKAMIAER 158
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
119 MEVFEIKERRGVARFPGVAKVDGEVACEAEM--MCARR 154

RESULT 15

```

```

US-10-282-122A-68573
; Sequence 68573, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangau
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68573
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (3)..(3)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (12)..(12)
; OTHER INFORMATION: X=any amino acid
US-10-282-122A-68573

Query Match 41.5%; Score 344.5; DB 4; Length 172;
Best Local Similarity 45.6%; Pred. No. 3.4e-30;
Matches 72; Conservative 27; Mismatches 48; Indels 11; Gaps 4;

Qy 1 MEQSHQNLQSQPIEHITLQILPHRYPMLLVDRIELQANKKIYAYKNITNEDVFNQHEP 60
Db 25 MSDNH-TLQ---IEETLDLPHRYPLLVDRIEFGKFLRAVKNVSNPEPFFQGHFP 79

Qy 61 NKDIFPQVLIVGMAOTGGFLAFTSLMGFPDETAKTIVVFMTIDKVKFRIPVTPGDRLE 120
Db 80 GKXIFPQVLILEMAQATGILAKFSVGLKFP---GELYFFAALDGARFKRPVLPGDMV 135

Qy 121 YHLEVLKHGMIVQVGGTAQVDGKVAEAEIKAMIAER 158
Db 136 LEVEFIKERRGVARFKGAVKVDGEIACELEM--MCARR 171

```

RESULT 15

Search completed: January 10, 2006, 13:37:58
Job time : 61 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:32:15 ; Search time 9 Seconds
(without alignments)
150.054 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 830

Sequence: 1 MEQSHQNLQSQFFIEHLQI.....QVQKVVAAELKAMIAERD 159

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 61141 seqs, 8493638 residues

Total number of hits satisfying chosen parameters: 61141

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA_New.*

- 1: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 5: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
- 7: /cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	276	33.3	149	6	US-10-467-657-6110 Sequence 6110, Ap
2	276	33.3	149	6	US-10-467-657-7550 Sequence 7550, Ap
3	276	33.3	154	6	US-10-467-657-6130 Sequence 6130, Ap
4	75.5	9.1	121	6	US-10-454-437-322 Sequence 322, App
5	75.5	9.1	121	6	US-10-454-437-324 Sequence 324, App
6	74.5	9.0	485	6	US-10-485-517-136 Sequence 136, App
7	73.5	8.9	393	7	US-11-194-246-309 Sequence 309, App
8	71.5	8.6	564	6	US-10-821-234-1340 Sequence 1340, Ap
9	70.5	8.5	912	7	US-11-042-988-12 Sequence 12, Appl
10	66	8.0	205	7	US-11-093-746A-26 Sequence 26, Appl
11	66	8.0	792	7	US-11-103-957-92 Sequence 92, Appl
12	65.5	7.9	129	7	US-11-052-554A-363 Sequence 363, App
13	65.5	7.9	323	7	US-11-055-822-628 Sequence 628, App
14	65	7.8	1343	7	US-11-052-554A-284 Sequence 284, App
15	64.5	7.8	463	6	US-10-510-386-186 Sequence 186, App
16	64	7.7	585	6	US-10-821-234-1489 Sequence 1489, Ap
17	63.5	7.7	444	6	US-10-525-710-48 Sequence 48, Appl
18	63.5	7.7	559	6	US-10-873-528-158 Sequence 158, App
19	63.5	7.7	674	6	US-10-467-657-6812 Sequence 6812, Ap
20	63	7.6	332	6	US-10-793-626-550 Sequence 550, App
21	63	7.6	406	6	US-10-770-726-73 Sequence 73, Appl
22	62	7.5	514	6	US-10-840-688-2 Sequence 2, Appl
23	62	7.5	514	6	US-10-840-688-21 Sequence 21, Appl
24	62	7.5	792	6	US-10-467-657-6026 Sequence 6026, Ap
25	62	7.5	792	6	US-10-467-657-7528 Sequence 7528, Ap

26	62	7.5	887	6	US-10-467-657-3960 Sequence 3960, Ap
27	62	7.5	925	6	US-10-454-437-50 Sequence 50, Appl
28	61.5	7.4	481	7	US-11-090-439-16 Sequence 16, Appl
29	61	7.3	502	6	US-10-821-234-1554 Sequence 1554, Ap
30	61	7.3	514	6	US-10-840-688-3 Sequence 3, Appl
31	60.5	7.3	257	6	US-10-467-657-1522 Sequence 1522, Ap
32	60.5	7.3	950	6	US-10-467-657-854 Sequence 854, App
33	60.5	7.3	2340	7	US-11-052-554A-171 Sequence 171, App
34	60	7.2	514	6	US-10-840-688-10 Sequence 10, Appl
35	60	7.2	514	6	US-10-840-688-12 Sequence 12, Appl
36	60	7.2	323	6	US-10-467-657-2418 Sequence 2418, Ap
37	59.5	7.2	321	7	US-10-959-310-4 Sequence 4, Appl
38	59.5	7.2	321	7	US-11-131-212-72 Sequence 72, Appl
39	59.5	7.2	502	6	US-10-467-657-7332 Sequence 7332, Ap
40	59.5	7.2	1503	7	US-11-087-100-6 Sequence 6, Appl
41	59.5	7.2	1503	7	US-11-087-084-6 Sequence 6, Appl
42	59.5	7.2	1503	7	US-11-087-085-6 Sequence 6, Appl
43	59	7.1	263	7	US-11-082-389-120 Sequence 120, App
44	59	7.1	278	7	US-11-082-389-118 Sequence 118, App
45	59	7.1	338	7	US-11-016-584-11 Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-10-467-657-6110
; Sequence 6110, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6110
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6110

Query Match	33.3%	Score	276;	DB	6;	Length	149;
Best Local Similarity	41.5%	Pred. No.	5.1e-23;				
Matches	59;	Conservative	26;	Mismatches	51;	Indels	6;
Gaps	2;						
Qy	17	ILQLPHRYMLLVDRIRIELQANKKIVAYKNITNEDVFNHFPNKPFPVGLIVEGMAQ	76				
Db	12	IQKLPHRYFLQDRITAFEPKMTLTAIKNVITNEPQFGHFDLPVMPGVLLIEAMAQ	71				
Qy	77	TGGFLATSLWGPDPPEITAKTVVYFMTIDKVKFRIPVTPGDRLEVHLVXKHGMWQVG	136				
Db	72	ACGTLATLSEGG-----RKNEFFFFAGIDEARFKRQVIGDOLVFVELLTSRRGIGKFN	127				
Qy	137	GTAQVDGKVVAAELKAMIAER	158				
Db	128	AVAKVDGQVAEVI--MCAKR	147				

RESULT 2
US-10-467-657-7550
; Sequence 7550, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita

```

; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7550
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7550

```

Query Match	33.3%	Score	276;	DB	6;	Length	149;
Best Local Similarity	41.5%;	Pred. No.	5.1e-23;				
Matches	59;	Conservative	26;	Mismatches	51;	Indels	6;
Gaps	2;						
Qy	17	ILQILPHRYPMLLVDRILIELQANKKIYAYKNITINEDVFNCHGFNPKPIPGCVLIVEGMAQ	76				
Db	12	IQKLIPHRYPFLQDRITAFEPMTKLTAIKKNVTINEPQFGHFDDLPMVPGVLLIEAWAQ	71				
Qy	77	TGGFLAFTSLWGDFDEIATKTKIVYFMITDKVKFRIPVTPGDRLEYHLVLEHKGMVQVG	136				
Db	72	ACGTALISLGG----RKENEFFFPFAGIDEARFKRQVIPGDQLVFEVELLTSRRGICKFN	127				
Qy	137	GTAQVDGKVVAAELKAMIAER	158				
Db	128	AVAKVDGQVAEVI--MCAKR	147				

```

RESULT 3
US-10-467-657-6130
; Sequence 6130, Application US/10467657
; Publication No. US20050260581a1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEIN
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6130
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6130

```

Query Match	33.3%;	Score	276;	DB	6;	Length	154;
Best Local Similarity	41.5%;	Pred. No.	5.4e-23;				
Matches	59;	Conservative	26;	Mismatches	51;	Indels	6;
Gaps	2;						

Qy	17	ILQLPHRYPMLLVDRILIELQANKKIIVAYKNITNEDEVNGHGFENPKPIFGVLIVEGMAQ	76
Db	17	IQKLIHPHYPPQLQDRITAFEPMTLTAIKNVTINPEQFGHFPDLVPMPGVLIIEMAAQ	76
Qy	77	TGGFLAFTSLWGDFPEIAKTQIVYFMTIDKVKFRIPVTPGDRLEHYHLVHLKHGMIVQVG	136
Db	77	ACGTALILSEGG----RKENEFFFPAGIDEARFKRQVTPGDLQVFEVELLTSRRGIGKFN	132
Qy	137	GTAQVDGKVAEAEIKAMIAER	158
Db	133	AVAKVDGQVAEAVI--MCAKR	152

```

RESULT 4
US-10-454-437-322
; Sequence 322, Application US/10454437
; Publication No. US20050277115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CPCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 322
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-322

Query Match          9.1%; Score 75.5; DB 6; Length 121;
Best Local Similarity 28.7%; Pred. No. 0.17;
Matches 27; Conservative 15; Mismatches 35; Indels 17; Gaps

Qy 73 GMAQTGGFLAFTSL---WG--FDEIATKIVYFWTIDKVKFRIPVTPGDRLEY-----
Db 27 GGAIAHGFLTSMIIPFWGELLDTGVTTKVY--GLDKVRFSPVKVGSIRMGAVVR

Qy 123 LEVLKHGKMTWGGTQAQVDGKVAEAEKAMIA 156
Db 85 ISEYKGNGLHLVADGTIEEGQ-----ERPAAVA 113

RESULT 5
US-10-454-437-324
; Sequence 324, Application US/10454437
; Publication No. US20050277115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CPCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13

```


; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 324
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-324

Query Match 9.1%; Score 75.5; DB 6; Length 121;
Best Local Similarity 28.7%; Pred. No. 0.17;
Matches 27; Conservative 15; Mismatches 35; Indels 17; Gaps 5;
QY 73 GMAQTGGFLAFTSL---WG--FDPEIAKTKIVYFMTIDKVKRIPVTPGDRLEY-----H 122
Db 27 GGAIAHGFLTSLMIIPFWGELLDVTGTTKVNY--GLDKVRFTSPVKVGSRIKMGAVRE 84
QY 123 LEVLKHKMIMQVGGTAQVDGKVAEAEKAMIA 156
Db 85 ISEVKGNGLHLVADGTITIEGQ-----ERPAVVA 113

RESULT 6
US-10-485-517-136
; Sequence 136, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynex Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629W0
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 136
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-136

Query Match 9.0%; Score 74.5; DB 6; Length 485;
Best Local Similarity 23.8%; Pred. No. 1.3;
Matches 31; Conservative 21; Mismatches 47; Indels 31; Gaps 6;
QY 2 ESHONLQSQPFIEHILQILPHRYPMLLDVRIELQ-----ANKKIVAYKNITFNEVDVF 55
Db 249 BEQYGLNEFFFCG---EILPHRAKALVNRPPSATVNTYGTTEATVATVTSIQITQEIL 305

QY 56 NQHPPNKPI-----PPGVLIVEGMAQTGGFL-----AFTSLWGFDPPIAKTKIV 99
Db 306 D-QYPTLVGVVERPGCARLSTTDEGBLVIEGQSVSLGYLKNQDKQTAEVNFDDGIRT---- 360
QY 100 YFMTIDKVKF 109
Db 361 -YHTGDKAKF 369

RESULT 7
US-11-194-246-309
; Sequence 309, Application US/11194246
; Publication No. US20050272089A1
; GENERAL INFORMATION:
; APPLICANT: Mott, John
; APPLICANT: Trepod, Catherine
; APPLICANT: Arvidson, Staffan
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND METI
; TITLE OF INVENTION: USE
; FILE REFERENCE: 00592.US1 (M&R 268.05920101)
; CURRENT APPLICATION NUMBER: US/11/194,246
; CURRENT FILING DATE: 2005-08-01
; PRIOR APPLICATION NUMBER: US/10/274,586
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/345,438
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 621
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 309
; LENGTH: 393
; TYPE: PRT
; ORGANISM: HAEMOPHILUS INFLUENZAE
US-11-194-246-309

Query Match 8.9%; Score 73.5; DB 7; Length 393;
Best Local Similarity 21.6%; Pred. No. 1.3;
Matches 44; Conservative 24; Mismatches 47; Indels 89; Gaps 10;
QY 16 HILQLPHRYPMLLDVRIELQANKKIVAYKNITFNEVDVFNHGFNFKPIFGVLIVEGM- 74
Db 189 HIIRDLPPEY-----KIYSEKNFTFNK--IKQANENGILLWDKTIINVDGMK 231
QY 75 -----AQTGGFLAFTSL-----WGFI-----D 90
Db 232 TGHTSQAGYNLVASATTSNNRLLISVVMGVPTKYGREVESKKLLQWGFANFETFKLEAG 291
QY 91 PEIAKTKIVY-----FMTIDK-----VKPRIPVTPGDRLEYHLEVLKHKGM 131
Db 292 KEISEQRVYVYGDKNISVKLGALMDHFITIPKQKQSEVKARYELA-----DKNLQAPLVKQ 346
QY 132 IWQVGGTA-QVDGKVAEAEKAM 154
Db 347 V--IGKVVYQLDGKDIIASANLQVM 368

RESULT 8
US-10-821-234-1340
; Sequence 1340, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Presclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0

```

; SEQ ID NO 1340
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1340

Query Match      8.6%; Score 71.5; DB 6; Length 564;
Best Local Similarity 21.7%; Pred. No. 3.4;
Matches 39; Conservative 27; Mismatches 59; Indels 55; Gaps 9;

QY 2 ESHQNLQSQPFIEHILQILPHRYPMMLVDRIELQANKK-----IV 43
DB 365 EQMHA-----LLAALTYPM-RIDESIHLQREKYGDKMLRMQKQDPQVVEELF 413

QY 44 AYKNITF-----NEDEVFNHPPNKPPIPGVLI-----VEGMAQTGGFLAFTSLWGFDPRI 93
DB 414 SYSCPKFLSPVVPVNDVHPNKHFFLQQLKVFDEVOQQQLSTIRFLKLYTMP-- 471

QY 94 AKTKIVYFMTIDKVKFRIPVTPGDRLEHVLKHKM--KGMIVQVGTAQVDGKVAEAE 151
DB 472 -VAKLAGFLDLTEQEFRI-----QLLVFKHKMKNLVW-TSGISALDGEFQSASEV 519

RESULT 9
US-11-042-988-12
; Sequence 12, Application US/11042988
; Publication No. US20050244818A1
; GENERAL INFORMATION:
; APPLICANT: SILICIANO, ROBERT
; APPLICANT: ZHANG, HAILI
; APPLICANT: ZHOU, YAN
; TITLE OF INVENTION: SINGLE CELL ANALYSIS OF HIV REPLICATION CAPACITY AND
; FILE OF INVENTION: DRUG RESISTANCE
; FILE REFERENCE: 62760(71699)
; CURRENT APPLICATION NUMBER: US/11/042,988
; CURRENT FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: 60/540,716
; PRIOR FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent In Ver. 3.3
; SEQ ID NO 12
; LENGTH: 912
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-11-042-988-12

Query Match      8.5%; Score 70.5; DB 7; Length 912;
Best Local Similarity 23.6%; Pred. No. 8.1;
Matches 37; Conservative 26; Mismatches 45; Indels 49; Gaps 8;

QY 17 ILQILPHRYPMMLVDRIELQANKK-----IVAYKNITFNEDEVFNHPPNKPPIFG--- 67
DB 569 IIAQPDQSESELVNIQILKEKVKYLAWPVHAHGIGNEQV-----DKLVSAGIRK 622

QY 68 VLIVGMAQT-----GGFLAFTSLWGFDPRIAKTKIVYFMTIDKVKFR-----IP 112
DB 623 VFLDGDIDKAQDEHEKHENWRAMASDFNLPPVAKIIVA---SCDKCQLKGAMHCQVD 679

QY 113 VTPGDRLEHVLKHKMIVQVGTAQVDGKVAEAE 149
DB 680 CSPG-----IWQLDCT-HLEGGKVLVA 700

```

```

RESULT 10
US-11-093-746A-26
; Sequence 26, Application US/11093746A
; Publication No. US2005026643A1
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo M.
; APPLICANT: Callin, George A.
; TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE AND
; FILE OF INVENTION: COMPOSITIONS AND METHODS FOR MAKING AND USING THE SAME
; FILE REFERENCE: 3589.1015-003

```

```

; CURRENT APPLICATION NUMBER: US/11/093,746A
; CURRENT FILING DATE: 2005-03-30
; PRIOR APPLICATION NUMBER: PCT/US2003/032270
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/417,842
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-11-093-746A-26

Query Match      8.0%; Score 66; DB 7; Length 205;
Best Local Similarity 26.0%; Pred. No. 3.6;
Matches 13; Conservative 16; Mismatches 13; Indels 8; Gaps 2;

QY 95 KTKIVYFMTIDKVKFRIPVTPGDRLEHVLKHKM---IWQVGGTAQV 141
DB 30 KTWLYKLKLGCVTVTPT-----IGFNLTVYKGINFTVWDIGGQEKI 74

RESULT 11
US-11-103-957-92
; Sequence 92, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent Georges Christian Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/11/103,957
; CURRENT FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92
; LENGTH: 792
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-11-103-957-92

Query Match      8.0%; Score 66; DB 7; Length 792;
Best Local Similarity 28.4%; Pred. No. 21;
Matches 27; Conservative 10; Mismatches 28; Indels 30; Gaps 5;

QY 79 GFLAFT-----SLWG---FDPEIARTKIVYFMTIDK---VKFRIPVTPGDRLE 120
DB 469 GSLSTDPYFTADGVSGLGYDIYKAFDPRKASTVKQYKTTAGGVRMGIPVTEYDRVN 528

QY 121 Y-----HLEV-----LKHKGMIWQVGTAQVDG 143
DB 529 FGLAAEHLTVNTYNKAPRYADFIKQYKTDGADG 563

RESULT 12
US-11-052-554A-363
; Sequence 363, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A

```

```

; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 363
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-363

Query Match
Best Local Similarity 7.9%; Score 65.5; DB 7; Length 129;
Matches 26; Conservative 18; Mismatches 39; Indels 19; Gaps 4;

Qy 33 IIELOANK--KIVAYKNITNEOVFNHGRNKPFPQVLVEGMAQ---TGGFLAFTSLW 87
Db 41 IVKTDNKTDKVAFRKATGVMVDLVKDKYKVKSGGFTTIDGVTQKKAGRY-----W 94

Qy 88 GFDPEIAKTKIVYFMTIDKVKFRIPVTPGDRLEHYHLEVLKHK 129
Db 95 MFD-----VNDKLASKAADKIKVNGDKIEFYLVKVKYK 128

RESULT 13
US-11-055-822-628
; Sequence 628, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CPCN
; CURRENT APPLICATION NUMBER: US/11/055,822
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 628
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-628

Query Match
Best Local Similarity 7.9%; Score 65.5; DB 7; Length 323;
Matches 28; Conservative 18; Mismatches 38; Indels 35; Gaps 6;

Qy 48 ITFNEVFNCHFP-NKPIF-----PGVLIVEGM--AQTGGFLAFTSLW 87
Db 170 LRFVTDVKSGKLEVNAPVSYHTAYDRVPCEFTTWQPDILIVEGLNVLQTGPTLWVSDLF 229

Qy 88 GFD-----PEIAKTKIVYFMTIDKVKFRIPVTPGDRLEHYHLEVLKHKGM-----IWQ 134
Db 230 DFSVVDARTEDIEKWYIDRFKLKLDTAFR---RFGAHFHYADMAPDSIAVARELWQ 285

RESULT 14
US-11-052-554A-284
; Sequence 284, Application US/11052554A
; Publication No. US20050288666A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 284
; LENGTH: 1343
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-284

Query Match
Best Local Similarity 7.8%; Score 65; DB 7; Length 1343;
Matches 33; Conservative 10; Mismatches 37; Indels 26; Gaps 7;

Qy 67 GVLIVEGMAQT---GGFLAFTSLWGFDPPEIAKT-KIYFMTI---DKVKFRIPVTPGDR 118
Db 924 GVQLVK--AQTIVIGSGSLTLDLQGDVEADKTLHLIAQNGTVAEGDYGFRLTTAPGDG 981

Qy 119 L--BYHLEVL-----KHGMIWQVGGTAQVDGKVVABEL 151
Db 982 LYNVYGLKALNIHGQKLTAEHGG---AYGATADMSAKIGGDL 1024

RESULT 15
US-10-510-386-186
; Sequence 186, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 186
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-186

Query Match
Best Local Similarity 7.8%; Score 64.5; DB 6; Length 463;
Matches 22; Conservative 14; Mismatches 26; Indels 33; Gaps 5;

Qy 64 IFFGVLIIV-----EGMAQTGGFL-----AFTSLWGFDPPEIAKTIVYFM 102
```

Db 49 IIGGILVILIGLYAELSSAIPETGGGLIFVYRAFGRKTAFAAAG-----VLFG 98

Qy 103 TIDKVKFRIPVTPGDRLEYHLEVLKXHKGMWQVGG 137

Db 99 YVSVITFEAVALP-TVIDYVLPV-EHQGFMSLSG 131

Search completed: January 10, 2006, 13:38:18
Job time : 9 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:35:05 ; Search time 77 Seconds
(without alignments)
907.289 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 159

Sequence: 1 MQSHQNLSQFIEHLQI.....QVDGKVVAELKAMIAERD 159

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2443163 seqs, 439378781 residues

Word size : 8

Total number of hits satisfying chosen parameters: 40

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : A_Geneseq_21.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*
- 9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	159	100.0	159	6	ABU31058 Protein e
2	118	74.2	159	4	Aau36008 Halicobac
3	12	7.5	139	6	Abu46781 Protein e
4	12	7.5	140	3	Aab03793 Fatty aci
5	12	7.5	140	4	Aau38023 Streptoco
6	12	7.5	140	4	Aam01011 CFE 10 pr
7	12	7.5	140	5	Abp28016 Streptoco
8	12	7.5	140	5	Abp28015 Streptoco
9	12	7.5	140	6	Abu00778 S. pneumo
10	12	7.5	140	6	Abu45885 Protein e
11	12	7.5	140	8	Adk47391 Streptoco
12	12	7.5	140	8	Adv89450 Streptoco
13	12	7.5	140	8	Adv80703 Streptoco
14	12	7.5	140	8	Adv82870 Streptoco
15	12	7.5	146	8	Adr94257 Novel S.
16	12	7.5	146	9	Aea58127 Streptoco
17	12	7.5	160	4	Aae02200 Streptoco
18	11	6.9	141	6	Abu23918 Protein e
19	10	6.3	146	6	Abu26345 Protein e
20	9	5.7	102	6	Abu27427 Protein e
21	9	5.7	102	6	Abu27429 Protein e
22	9	5.7	146	9	Aea49209 L. rhammo
23	9	5.7	151	4	Aau34452 E. coli c
24	9	5.7	151	6	Abu15025 Protein e

25	9	5.7	151	6	ABU23117	Abu23117 Protein e
26	9	5.7	151	8	ADN18054	Adn18054 Bacterial
27	9	5.7	157	8	ADN17554	Adn17554 Bacterial
28	9	5.7	162	7	ADF06342	Adf06342 Bacterial
29	9	5.7	172	6	ABU40649	Abu40649 Protein e
30	9	5.7	172	8	ADS42775	Ads42775 Bacterial
31	9	5.7	181	6	ABU50043	Abu50043 Protein e
32	8	5.0	133	3	AAU90363	Aay90363 S. zooepl
33	8	5.0	137	8	ADN20318	Adn20318 Bacterial
34	8	5.0	139	6	ABU30001	Abu30001 Protein e
35	8	5.0	152	6	ABU39542	Abu39542 Protein e
36	8	5.0	154	8	ADS30417	Ads30417 Bacterial
37	8	5.0	161	3	AAU12580	Aag12580 Zea mays
38	8	5.0	437	8	ADX77098	Adx77098 Planc ful
39	8	5.0	478	8	ADN19962	Adn19962 Bacterial
40	8	5.0	910	8	ADS43079	Ads43079 Bacterial

ALIGNMENTS

RESULT 1

ABU31058	
ID	ABU31058 standard; protein; 159 AA.
XX	
AC	ABU31058;
XX	
DT	19-JUN-2003 (first entry)
XX	
DE	Protein encoded by Prokaryotic essential gene #16585.
XX	
KW	Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX	
OS	Helicobacter pylori.
XX	
PN	WO200277183-A2.
XX	
PD	03-OCT-2002.
XX	
PF	21-MAR-2002; 2002WO-US009107.
XX	
PR	21-MAR-2001; 2001US-00815242.
PR	06-SEP-2001; 2001US-00948993.
PR	25-OCT-2001; 2001US-0342923P.
PR	08-FEB-2002; 2002US-00072851.
PR	06-MAR-2002; 2002US-0362699P.
XX	
PA	(ELIT-) ELITRA PHARM INC.
XX	
PI	Wang L, Zamudio C, Malone C, Hasselbeck R, Ohlsen KL, Zyskind JW;
PI	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX	
DR	WPI; 2003-029926/02.
XX	
XX	N-PSDB; ACA34928.
PT	New antisense nucleic acids, useful for identifying proteins or screening
PT	for homologous nucleic acids required for cellular proliferation to
PT	isolate candidate molecules for rational drug discovery programs.
XX	
PS	Claim 25; SEQ ID NO 58982; 1766pp; English.
XX	

The invention relates to an isolated nucleic acid comprising any one of

of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 159 AA;

Query Match 100.0%; Score 159; DB 6; Length 159;
 Best Local Similarity 100.0%; Pred. No. 1.6e-161; Mismatches 0; Gaps 0;
 Matches 159; Conservative 0;

QY 1 MEQSHQNLQSQFFIEHILQILPHRYPMLLVDRILIELOANKKIVAYKNITFNEVDVFNHGFP 60
 DB 1 MEQSHQNLQSQFFIEHILQILPHRYPMLLVDRILIELOANKKIVAYKNITFNEVDVFNHGFP 60
 QY 61 NRPFPFGLVIEGMAQTGGFLAFTSLWGFDEPEIAKTIVYFMTIDKVKFRIPVTPGDRLE 120
 DB 61 NRPFPFGLVIEGMAQTGGFLAFTSLWGFDEPEIAKTIVYFMTIDKVKFRIPVTPGDRLE 120
 QY 121 YHLEVLKHKGMIVQVGTAQVDGKVAEAEELKAMIAERD 159
 DB 121 YHLEVLKHKGMIVQVGTAQVDGKVAEAEELKAMIAERD 159

RESULT 2
 AAU36008
 ID AAU36008 standard; protein; 159 AA.
 AC AAU36008;
 XX
 DT 14-FEB-2002 (first entry)

XX Helicobacter pylori cellular proliferation protein #321.

XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
 XX antibacterial; drug design.

XX Helicobacter pylori.

OS WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207272P.

PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.

PR 22-DEC-2000; 2000US-0257931P.

PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX

DR WPI; 2001-611495/70.
 XX N-PSDB; AAS53867.

XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.

XX Example 3; SEQ ID NO 11601; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
 CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
 CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 159 AA;

Query Match 74.2%; Score 118; DB 4; Length 159;

Best Local Similarity 100.0%; Pred. No. 1.2e-117;
 Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 KIVAYKNITFNEVDVFNHGFPKPIPGVLIVEGMAQTGGFLAFTSLWGFDEPEIAKTIVY 100
 DB 41 KIVAYKNITFNEVDVFNHGFPKPIPGVLIVEGMAQTGGFLAFTSLWGFDEPEIAKTIVY 100
 QY 101 FMTIDKVKFRIPVTPGDRLEHVLKHKGMIVQVGTAQVDGKVAEAEELKAMIAER 158
 DB 101 FMTIDKVKFRIPVTPGDRLEHVLKHKGMIVQVGTAQVDGKVAEAEELKAMIAER 158

RESULT 3
 ABU46781
 ID ABU46781 standard; protein; 139 AA.

XX AC ABU46781;

XX 19-JUN-2003 (first entry)

XX Protein encoded by prokaryotic essential gene #32308.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Streptococcus pyogenes.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX

DR WPI; 2003-029926/02.
 DR N-PSDB; ACA50651.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 25; SEQ ID NO 74705; 1766pp; English.
 PS
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 139 AA;
 SQ

Query Match 7.5%; Score 12; DB 6; Length 139;
 Best Local Similarity 100.0%; Pred. No. 0.0003;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
 |||||
 DB 10 LPHRYPMLLVDR 21

RESULT 4
 AAB03793
 ID AAB03793 standard; protein; 140 AA.
 XX
 XX AAB03793;
 XX
 DT 13-OCT-2000 (first entry)
 XX
 DE Fatty acid biosynthetic pathway protein FabZ.
 XX
 XX Fatty acid biosynthesis; FabZ; beta-hydroxyacyl-ACP; trans-2-acyl-ACP;
 KW fatty acid elongation; Streptococcus pneumoniae; antibody.
 KW
 XX Streptococcus pneumoniae.
 OS
 XX WO200030662-A1.
 PN
 XX 02-JUN-2000.
 PD
 XX 09-NOV-1999; 99WO-US026435.
 PF
 XX

PR 19-NOV-1998; 98US-00196388.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA
 XX Warren PV, Konstantinidis A, Russell RB;
 PI
 XX WPI; 2000-548599/50.
 DR N-PSDB; AAAS9854.
 DR
 XX Streptococcus pneumoniae fabZ proteins useful for diagnosing and treating
 PT microbial infections.
 PT
 XX Claim 1; Page 3; 53pp; English.
 PS
 XX This sequence represents the fatty acid biosynthetic pathway protein
 CC FabZ. The invention relates to FabZ nucleotide and protein sequences.
 CC FabZ catalyzes the dehydration of nascent beta-hydroxyacyl-ACP to trans-2
 CC -acyl-ACP in the process of fatty acid elongation. The invention also
 CC includes an antibody specific for the FabZ nucleotide sequence, and a
 CC method for the treatment of an individual in need of enhanced FabZ
 CC expression. Also included is a method for diagnosing or prognosing a
 CC disease or a susceptibility to a disease in an individual related to the
 CC expression of or activity of a FabZ protein. FabZ nucleotide sequences
 CC and protein sequences may be used in the prevention, treatment and
 CC diagnosis of diseases associated with fabZ expression and Streptococcal
 CC infection. FabZ nucleotide and protein sequences may be used to treat
 CC diseases by rectifying defects in a genome that affect the activity of
 CC fabZ. FabZ nucleotide sequences can be used to produce FabZ protein, and
 CC also may be used as probes and primers in diagnostic assays. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC against fabZ and in assays to identify modulators of fabZ expression and
 CC activity. The anti-fabZ antibodies and fabZ antagonists may also be used
 CC to down regulate fabZ expression and activity. They may be used to treat
 CC *S. pneumoniae* infections. The anti-fabZ antibodies may also be used as
 CC diagnostic agents for detecting the presence of fabZ polypeptides in
 CC samples
 XX
 SQ Sequence 140 AA;
 Query Match 7.5%; Score 12; DB 3; Length 140;
 Best Local Similarity 100.0%; Pred. No. 0.0003;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
 |||||
 DB 11 LPHRYPMLLVDR 22

RESULT 5
 AAU38023
 ID AAU38023 standard; protein; 140 AA.
 XX
 XX AAU38023;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Streptococcus pneumoniae cellular proliferation protein #452.
 XX
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200170955-A2.
 XX
 XX 27-SEP-2001.
 PD
 XX 21-MAR-2001; 2001WO-US009180.
 PF
 XX 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR

PI Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;
 PI Tettelin H;
 XX WPI; 2002-352536/38.
 DR N-PSDB; ABN68647.
 XX New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.
 XX Claim 1; Page 3684; 4525pp; English.
 PS The invention relates to a protein (ABP25413-ABP30895) from group B
 CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC nucleic acids encoding (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX Sequence 140 AA;
 SQ
 Query Match 7.5%; Score 12; DB 5; Length 140;
 Best Local Similarity 100.0%; Pred. No. 0.00031;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 LPHRYPMLLVDR 32
 Db 11 LPHRYPMLLVDR 22
 RESULT 8
 ABP28015
 ID ABP28015 standard; protein; 140 AA.
 AC ABP28015;
 XX 02-JUL-2002 (first entry).
 DE Streptococcus polypeptide SEQ ID NO 5206.
 DE Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW Group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX Streptococcus agalactiae.
 OS
 XX WO200234771-A2.
 PN 02-MAY-2002.
 XX 29-OCT-2001; 2001WO-GB004789.
 XX 27-OCT-2000; 2000GB-00026333.
 PR 24-NOV-2000; 2000GB-00028727.
 PR 07-MAR-2001; 2001GB-00005640.
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;
 PI Tettelin H;

DR WPI; 2002-352536/38.
 DR N-PSDB; ABN68646.
 XX New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.
 XX Claim 1; Page 3684; 4525pp; English.
 PS The invention relates to a protein (ABP25413-ABP30895) from group B
 CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC nucleic acids encoding (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX Sequence 140 AA;
 SQ
 Query Match 7.5%; Score 12; DB 5; Length 140;
 Best Local Similarity 100.0%; Pred. No. 0.00031;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 LPHRYPMLLVDR 32
 Db 11 LPHRYPMLLVDR 22
 RESULT 9
 ABU00778
 ID ABU00778 standard; protein; 140 AA.
 AC ABU00778;
 XX 23-OCT-2003 (revised)
 DT 11-FEB-2003 (first entry)
 XX S. pneumoniae type 4 strain protein from coding region #346.
 DE Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
 KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
 KW gene therapy; vaccine.
 XX Streptococcus pneumoniae; type 4 strain.
 OS
 XX WO200277021-A2.
 PN 03-OCT-2002.
 XX 27-MAR-2002; 2002WO-IB002163.
 XX 27-MAR-2001; 2001GB-00007658.
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX Massignani V, Tettelin H, Fraser C;
 XX WPI; 2003-040579/03.
 DR N-PSDB; ABX06056.
 XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT useful as medicaments for treating or preventing a disease or infection

PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
 XX ear infection.
 PS Claim 1; SEQ ID NO 692; 56pp; English.
 XX The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 of 2489 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC AB556454. Also included are an antibody which binds one of the proteins,
 CC treating a patient by administering the protein, DNA or antibody (in a
 CC composition), a kit comprising first and second primers, which are the
 CC nucleic acid cited above or fragments between nucleotides 8-100 of a
 CC sequence not defined in the specification, for amplifying a target
 CC sequence contained within a Streptococcus nucleic acid sequence, where
 CC the first primer is substantially complementary to the target sequence of
 CC the target sequence, and where the parts of the primers having
 CC substantial complementarity define the termini of the target sequence to
 CC be amplified, assay comprising contacting a test compound with the
 CC protein, and determining whether the test compound binds to the protein
 CC and a Streptococcus pneumoniae bacterium, where one or more genes
 CC encoding the proteins has been rendered inactive. The proteins, nucleic
 CC acid molecules, antibody and compositions are useful as medicaments for
 CC treating or preventing a disease or infection due to streptococcus
 CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
 CC media or ear infection. They are also useful in developing vaccines,
 CC diagnostics and antibiotics. The methods are useful for identifying
 CC immunodominant proteins. The present sequence is one of the 2469 proteins
 CC expressed by the identified coding regions from the genomic sequence.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
 CC standardise OS field)
 XX Sequence 140 AA;
 SQ

Query Match 7.5%; Score 12; DB 6; Length 140;
 Best Local Similarity 100.0%; Pred. No. 0.00031;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMMLVDR 32
 DB 11 LPHRYPMMLVDR 22
 |||||
 |||||

RESULT 10
 ABU45885
 ID ABU45885 standard; protein; 140 AA.
 XX
 AC ABU45885;
 XX
 XX 19-JUN-2003 (first entry)
 DT
 DE Protein encoded by Prokaryotic essential gene #31412.
 XX
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Streptococcus pneumoniae.
 XX
 OS WO200277183-A2.
 XX
 PN 03-OCT-2002.
 XX
 XX 21-MAR-2002; 2002WO-US009107.
 PF
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 PR

PA (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI: 2003-029926/02.
 DR N-PSDB; ACA49755.
 DR
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 73809; 1766pp; English.
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than S. aureus, S. typhimurium,
 CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 140 AA;

Query Match 7.5%; Score 12; DB 6; Length 140;
 Best Local Similarity 100.0%; Pred. No. 0.00031;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMMLVDR 32
 DB 11 LPHRYPMMLVDR 22
 |||||
 |||||

RESULT 11
 ADK47391
 ID ADK47391 standard; protein; 140 AA.
 XX
 AC ADK47391;
 XX
 XX 20-MAY-2004 (first entry)
 DT
 DE Streptococcus pneumoniae protein, Seq ID No 3906.
 XX
 XX Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
 KW Streptococcus pneumoniae.
 XX
 OS US6699703-B1.
 XX
 PN

PD 02-MAR-2004.
XX 26-MAY-2000, 2000US-00583110.
XX
PR 02-JUL-1997; 97US-0051553P.
PR 12-MAY-1998; 98US-0085131P.
PR 30-JUN-1998; 98US-00107433.
XX (GENO-) GENOME THERAPEUTICS CORP.
PA Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CB;
PI WPI; 2004-212399/20.
DR N-PSDB; ADR44730.
XX
XX New nucleic acid molecules and polypeptides useful for diagnosing,
PT preventing and treating pathological conditions resulting from bacterial
PT infection, e.g. Streptococcus pneumoniae infection, and in drug
PT screening.
XX
XX Disclosure; SEQ ID NO 3906; 301pp; English.
XX
XX The invention relates to isolated Streptococcus pneumoniae nucleic acids
CC and polypeptides. The nucleic acids and proteins are useful for
CC diagnosing, preventing and treating pathological conditions resulting
CC from bacterial infection, such as S. pneumoniae infection. These may also
CC be used for drug screening procedures. The present sequence represents a
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
CC data for this patent did not appear in the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
XX
SQ Sequence 140 AA;

Query Match 7.5%; Score 12; DB 8; Length 140;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
DB 11 LPHRYPMLLVDR 22

RESULT 12
ADV89450
ID ADV89450 standard; protein; 140 AA.
XX
AC ADV89450;
XX
XX 24-FEB-2005 (first entry)
XX
DE Streptococcus agalactiae protein sequence, SEQ ID 1844.
XX
KW Antibacterial; Vaccine; bacterial infection.
XX
OS Streptococcus agalactiae.
XX
PN FR2824074-A1.
XX
PD 31-OCT-2002.
XX
PF 26-APR-2001; 2001FR-00005642.
XX
PR 26-APR-2001; 2001FR-00005642.
XX
XX (INSP) INST PASTEUR.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Glaser P, Rusniok C, Chevalier F, Frangeul L, Lalioui L;
PI Zouine M, Couve E, Buchrieser C, Poyart C, Trieu CB, Kunst F;
XX WPI; 2004-101891/11.
XX

PT Genomic nucleotide sequences encoding polypeptides of Streptococcus
PT agalactiae for the development of vaccines, diagnostic tools, DNA chips
XX and identification of therapeutic targets.
XX
XX Claim 6; SEQ ID NO 1844; 2687pp; French.
XX
XX The present invention relates to novel Streptococcus agalactiae
CC nucleotide sequences (I; ADV87607-ADV87745) and novel polypeptides (II;
CC ADV87746-ADV89950). The nucleotide sequences encode polypeptides of S.
CC agalactiae involved in the synthesis of amino acids, cell membranes,
CC intermediate (central) metabolism, energetic metabolism, fatty acid and
CC phospholipid metabolism, nucleotide metabolism including purines,
CC pyrimidines and/or nucleosides, regulatory functions, replication,
CC transcription, translation, protein transport, adaptation to atypical
CC conditions, sensitivity to medicines and/or analogues, functions related
CC to transporters, biosynthesis of cofactors, prosthetic groups and
CC transporters, cell membrane proteins and cellular machinery. (I) are
CC useful for the detection and/or amplification of nucleic acids.
CC Pharmaceutical composition comprising (I) or (II) are useful for
CC treatment of a bacterial S. agalactiae infection. Note: WO200292818A2 is
CC equivalent for the present basic patent FR2824074A1. WO200292818A2
CC contains 6617 sequence whereas the present patent only contains 2344
XX sequences.
XX
SQ Sequence 140 AA;

Query Match 7.5%; Score 12; DB 8; Length 140;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
DB 11 LPHRYPMLLVDR 22

RESULT 13
ADV80703
ID ADV80703 standard; protein; 140 AA.
XX
AC ADV80703;
XX
DT 24-FEB-2005 (first entry)
XX
DE Streptococcus agalactiae protein, SEQ ID 1844.
XX
KW Antibacterial; vaccine; bacterial infection.
XX
OS Streptococcus agalactiae.
XX
PN WO200292818-A2.
XX
PD 21-NOV-2002.
XX
PF 26-APR-2002; 2002WO-IB003059.
XX
PR 26-APR-2001; 2001FR-00005642.
XX
XX (INSP) INST PASTEUR.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Glaser P, Rusniok C, Chevalier F, Frangeul L, Lalioui L;
PI Zouine M, Couve E, Buchrieser C, Poyart C, Trieu-Cuot P, Kunst F;
XX WPI; 2004-101891/11.
XX
XX Genomic nucleotide sequences encoding polypeptides of Streptococcus
PT agalactiae for the development of vaccines, diagnostic tools, DNA chips
PT and identification of therapeutic targets.
XX
XX Claim 6; SEQ ID NO 1844; 439pp; French.
XX
XX The present invention relates to novel Streptococcus agalactiae
CC nucleotide sequences (I; ADV78860-ADV78998 and ADV83341-ADV85476) and

CC novel polypeptides (II; ADV78999-ADV81203 and ADV81205-ADV83340). The
 CC nucleotide sequences encode polypeptides of *S. agalactiae* involved in the
 CC synthesis of amino acids, cell membranes, intermediate (central)
 CC metabolism, energetic metabolism, fatty acid and phospholipid metabolism,
 CC nucleotide metabolism including purines, pyrimidines and/or nucleosides,
 CC regulatory functions related to atypical conditions, sensitivity to medicines
 CC and/or analogues, functions related to transposons, biosynthesis of
 CC cofactors, prosthetic groups and transporters, cell membrane proteins and
 CC cellular machinery. (I) are useful for the detection and/or amplification
 CC of nucleic acids. Pharmaceutical composition comprising (I) or (II) are
 CC useful for treatment of a bacterial *S. agalactiae* infection. The complete
 CC genome of *Streptococcus agalactiae* is given in ADV81204. Note: The
 CC present patent is an equivalent for the basic patent FR2824074A1, which
 CC contains only 2344 sequences.

XX Sequence 140 AA;

Query Match 7.5%; Score 12; DB 8; Length 140;
 Best Local Similarity 100.0%; Pred. No. 0.00031;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
 |||||
 DB 11 LPHRYPMLLVDR 22

RESULT 14

ADV82870
 ID ADV82870 standard; protein; 140 AA.

AC ADV82870;

DT 24-FEB-2005 (first entry)

XX *Streptococcus agalactiae* protein, SEQ ID 4011.

DE Antibacterial; vaccine; bacterial infection.

KW *Streptococcus agalactiae*.

OS WO200292818-A2.

PN 21-NOV-2002.

PD 26-APR-2002; 2002WO-IB003059.

XX 26-APR-2001; 2001FR-00005642.

PR (INSP) INST PASTEUR.

PA (CNRS) CNRS CENT NAT RECH SCI.

XX Glaser P, Rusniok C, Chevalier F, Frangeul L, Lalioui L;
 PI Zouine M, Couve E, Buchrieser C, Poyart C, Trieu-Cuot P, Kunst P;
 XX WPI; 2004-101891/11.

DR Genomic nucleotide sequences encoding polypeptides of *Streptococcus*

XX *agalactiae* for the development of vaccines, diagnostic tools, DNA chips
 PT and identification of therapeutic targets.

XX Claim 6; SEQ ID NO 4011; 439pp; French.

XX The present invention relates to novel *Streptococcus agalactiae*
 CC nucleotide sequences (I; ADV78860-ADV78998 and ADV83341-ADV85476) and
 CC novel polypeptides (II; ADV78999-ADV81203 and ADV81205-ADV83340). The
 CC nucleotide sequences encode polypeptides of *S. agalactiae* involved in the
 CC synthesis of amino acids, cell membranes, intermediate (central)
 CC metabolism, energetic metabolism, fatty acid and phospholipid metabolism,
 CC nucleotide metabolism including purines, pyrimidines and/or nucleosides,
 CC regulatory functions, replication, transcription, translation, protein
 CC transport, adaptation to atypical conditions, sensitivity to medicines
 CC and/or analogues, functions related to transposons, biosynthesis of

CC cofactors, prosthetic groups and transporters, cell membrane proteins and
 CC cellular machinery. (I) are useful for the detection and/or amplification
 CC of nucleic acids. Pharmaceutical composition comprising (I) or (II) are
 CC useful for treatment of a bacterial *S. agalactiae* infection. The complete
 CC genome of *Streptococcus agalactiae* is given in ADV81204. Note: The
 CC present patent is an equivalent for the basic patent FR2824074A1, which
 CC contains only 2344 sequences.

XX Sequence 140 AA;

Query Match 7.5%; Score 12; DB 8; Length 140;
 Best Local Similarity 100.0%; Pred. No. 0.00031;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
 |||||
 DB 11 LPHRYPMLLVDR 22

RESULT 15

ADR94257
 ID ADR94257 standard; protein; 146 AA.

XX ADR94257;

DT 16-DEC-2004 (first entry)

XX Novel *S. pneumoniae* protein sequence, SEQ ID 2892.

XX Meningitis; bacteraemia; pneumonia; otitis media; vaccine;
 KW bacterial infection.

XX *Streptococcus pneumoniae*.

XX US6800744-B1.

XX 05-OCT-2004.

XX 30-JUN-1998; 98US-00107433.

XX 02-JUL-1997; 97US-0051553P.

PR 12-MAY-1998; 98US-0085131P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2004-697205/68.

DR N-PSDB; ADR91654.

XX New isolated nucleic acid encoding a *Streptococcus pneumoniae*
 PT polypeptide, useful for diagnosing, preventing and/or treating
 PT pathological conditions resulting from the bacterial infection.

XX Disclosure; SEQ ID NO 2892; 151pp; English.

XX The invention relates to an isolated nucleic acid comprising a sequence
 CC encoding a *Streptococcus pneumoniae* ADR91366polypeptide, or its
 CC fragments, with any of 9 fully defined sequences (appearing as ADR94308,
 CC ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642, ADR95682,
 CC ADR96079) or any of the fully defined sequences appearing as ADR91705,
 CC ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR92366, ADR92650 or
 CC ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide
 CC sequences, or at least 40, 60 or 300 consecutive nucleotides, which is
 CC hybridizable under high stringency conditions to the nucleotide sequence.
 CC The nucleic acids and proteins are chosen from 5206 disclosed sequences.
 CC Also included are a recombinant expression vector comprising the isolated
 CC nucleic acid cited above operably linked to a transcription regulatory
 CC element, a cell comprising the recombinant expression vector and a probe
 CC comprising at least 20 consecutive nucleotides of the nucleotide
 CC sequences as cited above. The methods and compositions of the present
 CC invention are useful for the diagnosis, prevention and/or treatment of
 CC pathological conditions resulting from bacterial infection by

CC Streptococcus pneumoniae e.g. pneumonia, bacteraemia, meningitis and
 CC otitis media. The present sequence is one of the 2603 disclosed S.
 CC pneumoniae protein sequences. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=6800744B1.
 XX
 SQ Sequence 146 AA;

Query Match 7.5%; Score 12; DB 8; Length 146;
 Best Local Similarity 100.0%; Pred. No. 0.00032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
 DB 17 LPHRYPMLLVDR 28
 |||||
 |||||

RESULT 16
 AEA58127
 ID AEA58127 standard; protein; 146 AA.
 AC AEA58127;
 DT 25-AUG-2005 (first entry)
 XX Streptococcus pneumoniae ORF amino acid sequence SEQ ID NO:2892.
 DE bacterial infection; Streptococcus pneumoniae infection; antibacterial;
 KW vaccine.
 KW Streptococcus pneumoniae.
 OS
 PN US2005136404-A1.
 XX
 PD 23-JUN-2005.
 XX
 PF 10-JUL-2003; 2003US-00617320.
 PR 02-JUL-1997; 97US-0051553P.
 PR 12-MAY-1998; 98US-0085131P.
 PR 30-JUN-1998; 98US-00107433.
 XX (DOUC/) DOUCETTE-STAMM L A.
 PA (BUSH/) BUSH D.
 XX Doucette-Stamm LA, Bush D;
 PI
 XX WPI; 2005-477576/48.
 DR N-PSDB; AEA55524.
 XX
 PT New isolated nucleic acid molecules and encoded polypeptides useful for
 PT diagnosing, preventing or treating bacterial infections, particularly
 PT Streptococcus pneumoniae infection.
 XX
 PS Claim 5; SEQ ID NO 2892; 144pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule for detecting,
 CC preventing or treating pathological conditions resulting from bacterial
 CC infection. The isolated nucleic acid comprises: (a) any of the 2603
 CC nucleotide sequences of AEA55236 to AEA57838; (b) a nucleotide sequence
 CC encoding a Streptococcus pneumoniae polypeptide comprising any of the
 CC 2603 amino acid sequences of AEA57839 to AEA60441; or (c) a nucleotide
 CC sequence of at least 8 nucleotides in length, where the sequence is
 CC hybridizable to a nucleic acid having any of the nucleotide sequences in
 CC (a). Also described: (1) a recombinant expression vector comprising the
 CC above nucleic acid operably linked to a transcription regulatory element;
 CC (2) a cell comprising the recombinant expression vector; (3) producing an
 CC S. pneumoniae polypeptide; (4) a probe comprising a nucleotide sequence
 CC consisting of at least 8 nucleotides of any of AEA55236 to AEA57838; (5)
 CC treating a subject for S. pneumoniae infection; (6) a recombinant or
 CC substantially pure preparation of an S. pneumoniae polypeptide or its
 CC fragment, where the polypeptide is selected from AEA57839 to AEA60441;

CC (7) a vaccine composition for preventing or treating an S. pneumoniae
 CC infection, comprising an amount of the above nucleic acid or polypeptide;
 CC (8) detecting the presence of a Streptococcus nucleic acid in a sample;
 CC (9) a computer readable medium having recorded the nucleotide sequences
 CC of AEA55236 to AEA57838; (10) a computer based system for identifying
 CC fragments of the Streptococcus genome of commercial importance. The
 CC composition and methods are useful for diagnosing, preventing or treating
 CC bacterial infections, particularly S. pneumoniae infection. The present
 CC sequence represents a S. pneumoniae ORF amino acid sequence from the
 CC present invention. Note - The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from the USPTO web site.
 XX
 SQ Sequence 146 AA;

Query Match 7.5%; Score 12; DB 9; Length 146;
 Best Local Similarity 100.0%; Pred. No. 0.00032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
 DB 17 LPHRYPMLLVDR 28
 |||||
 |||||

RESULT 17
 AAE02200
 ID AAE02200 standard; protein; 160 AA.
 AC AAE02200;
 XX
 DT 11-SEP-2003 (revised)
 DT 31-JUL-2001 (first entry)
 XX Streptococcus pneumoniae His6-beta-hydroxyacyl-ACP dehydrase (FabZ).
 DE
 XX His6-FabZ; high throughput method; fatty acid biosynthesis; therapy;
 KW bacterial enzyme; biological agent screening; otitis media; empyema;
 KW bacterial tracheitis; acute epiglottitis; thyroiditis; lung abscess;
 KW infective endocarditis; secretory diarrhoea; splenic; retroperitoneal;
 KW intrarenal; perinephric; cerebral; cutaneous; abscess; blepharitis;
 KW conjunctivitis; keratitis; endophthalmitis; preseptal; orbital;
 KW cellulitis; dacryocystitis; epididymitis; abscess; toxic shock syndrome;
 KW impetigo; folliculitis; wound infection; bacterial myositis;
 KW septic arthritis; osteomyelitis; beta-hydroxyacyl-ACP dehydrase;
 KW acyl carrier protein.
 XX
 OS Streptococcus pneumoniae; 0100993.
 PN
 XX WO200130988-A1.
 XX
 PD 03-MAY-2001.
 XX
 PF 26-OCT-2000; 2000WO-US029451.
 XX
 PR 27-OCT-1999; 99US-0161775P.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Dewolf W, Kallender H, Lonadale JT;
 DR WPI; 2001-316332/33.
 DR N-PSDB; AAD06211.
 XX High throughput method for screening for biological agents against fatty
 PT acid biosynthesis comprises contacting a bacterial enzymatic pathway with
 PT enzymes e.g. malonyl-CoA:ACP transacylase.
 XX
 PS Claim 1; Page 13; 94pp; English.
 XX The present invention relates to a high throughput method for screening
 CC biological agents affecting fatty acid biosynthesis, comprises contacting
 CC a bacterial enzymatic pathway with enzymes. The method is used for

screening biological agents affecting fatty acid biosynthesis. Agonists and antagonists of fab (fatty acid biosynthesis) are used to inhibit, prevent or treat diseases such as infections of the upper respiratory tract (e.g. otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower respiratory (e.g. empyema, lung abscess), cardiac (e.g. infective endocarditis), gastrointestinal (e.g. secretory diarrhoea, splenic abscess, retroperitoneal abscess), central nervous system (e.g. cerebral abscess), eye (e.g. blepharitis, conjunctivitis, keratitis, endophthalmitis, preseptal and orbital cellulitis, dacryocystitis), kidney and urinary tract (e.g. epididymitis, intrarenal and perinephric abscess, toxic shock syndrome), skin (e.g. impetigo, folliculitis, cutaneous abscess, cellulitis, wound infection, bacterial myositis) and bone and joint (e.g. septic arthritis, osteomyelitis). The present sequence is *Streptococcus pneumoniae* 010093 His6-beta-hydroxyl-ACP (acyl carrier protein) dehydrase (Fab2). In fatty acid biosynthetic pathway, dehydration by Fab2 leads to trans-2-enoyl-ACP which is in turn converted to acyl-ACP by enoyl-ACP reductase (FabI). (Updated on 11-SEP-2003 to standardise OS field)

XX Sequence 160 AA;

Query Match 7.5%; Score 12; DB 4; Length 160;
 Best Local Similarity 100.0%; Pred. No. 0.00035; Mismatches 0; Indels 0; Gaps 0;
 Matches 12; Conservative 0;

QY 21 LPHRYPMMLVDR 32
 |||||
 DB 31 LPHRYPMMLVDR 42

RESULT 18
 ABU23918
 ID ABU23918 standard; protein; 141 AA.
 XX
 AC ABU23918;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #9445.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Clostridium acetobutylicum.
 XX
 PN WO200271783-A2.
 XX
 PD 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.
 DR N-PSDB; ACA27788.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 51842; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 141 AA;

Query Match 6.9%; Score 11; DB 6; Length 141;
 Best Local Similarity 100.0%; Pred. No. 0.0036;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 PHRYPMMLVDR 32
 |||||
 DB 13 PHRYPMMLVDR 23

RESULT 19
 ABU26345
 ID ABU26345 standard; protein; 146 AA.

XX
 AC ABU26345;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #11872.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Campylobacter jejuni.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.
 DR N-PSDB; ACA30215.

XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX XX
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX XX
XX PA (ELIT-) ELITRA PHARM INC.
XX XX
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX XX
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA31299.
XX XX
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX XX
XX PS Claim 25; SEQ ID NO 55353; 1766pp; English.
XX XX
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX XX
XX SQ Sequence 102 AA;
Query Match 5.7%; Score 9; DB 6; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.38; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 62 KPFPFGVLI 70
Db 74 KPFPFGVLI 82
|||||
|||||
RESULT 22
AEA49209
IP AEA49209 standard; protein; 146 AA.

XX AC AEA49209;
XX DT 25-AUG-2005 (first entry)
XX DE L. rhamnosus polypeptide #73.
XX KW Gene fusion; wound healing; immunostimulant; antilipemic; vulnerary.
XX OS Lactobacillus rhamnosus.
XX PN WO2005056801-A1.
XX PD 23-JUN-2005.
XX PF 15-DEC-2003; 2003WO-NZ000278.
XX PR 15-DEC-2003; 2003WO-NZ000278.
XX PA (GENE-) GENESIS RES & DEV CORP LTD.
XX PA (FONT-) FONTEREA COOP GROUP LTD.
XX PI Glenn M, Havukkala I, Lubbers MW, Dekker J;
XX DR WPI; 2005-435589/44.
XX DR N-PSDB; AEA49117.
XX XX
XX PT Isolated Lactobacillus polynucleotide, useful for modulating activity of
XX PT polypeptide in a cell, and for production of more effective probiotic
XX PT bacteria.
XX PS Claim 18; SEQ ID NO 153; 323pp; English.
XX XX
XX CC The invention relates to an isolated Lactobacillus rhamnosus
XX CC polynucleotide and the polypeptide it encodes. The invention also relates
XX CC to a genetic construct comprising the polynucleotide, a transgenic host
XX CC cell comprising the genetic construct, a transgenic organism comprising
XX CC the host cell or its progeny and a method of modulating the activity of a
XX CC polypeptide in an organism comprising stably incorporating the
XX CC polynucleotide into the genome of the organism. The polynucleotide is
XX CC useful for modulating the activity of a polypeptide in an organism, for
XX CC creating a transgenic organism and for the detection of lactic acid
XX CC bacteria in a sample. The polynucleotides and polypeptides may be used in
XX CC the selection and production of more effective probiotic bacteria, for
XX CC immune function enhancement, for reduction of blood lipids such as
XX CC cholesterol, for wound healing and in vaccine development. They are also
XX CC useful as health supplements, as animal probiotics for improved animal
XX CC health and for the production of improved food bacteria through genetic
XX CC modification. The polynucleotide may also be useful in genome mapping and
XX CC physical mapping. This sequence represents an *L. rhamnosus* polypeptide of
XX CC the invention.
XX SQ Sequence 146 AA;
Query Match 5.7%; Score 9; DB 9; Length 146;
Best Local Similarity 100.0%; Pred. No. 0.52; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 ILPHRYPML 28
Db 16 ILPHRYPML 24
|||||
|||||
RESULT 23
AAU34452
ID AAU34452 standard; protein; 151 AA.
XX AC AAU34452;
XX XX
XX DT 14-FEB-2002 (first entry)
XX DE E. coli cellular proliferation protein #33.
XX XX

KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
OS Escherichia coli.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.
XX 21-MAR-2000; 2000US-0191078P.
XX 23-MAY-2000; 2000US-0206848P.
XX 26-MAY-2000; 2000US-0207272P.
XX 23-OCT-2000; 2000US-0242578P.
XX 27-NOV-2000; 2000US-0253625P.
XX 22-DEC-2000; 2000US-0257931P.
XX 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlseen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
XX N-PSDB; AAS52311.
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.
XX Example 3; SEQ ID NO 10045; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
XX prokaryotic cellular proliferation, their use in identifying the genes,
XX their use in the discovery of novel antibiotics, the essential genes,
XX themselves and the encoded proteins. The prokaryotes used are Escherichia
XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX useful for the identification of potential new targets for antibiotic
XX development. The antisense nucleic acids can also be used to identify
XX proteins used in proliferation, to express these proteins, and to obtain
XX antibodies capable of binding to the expressed proteins. The proteins can
XX be used to screen compounds in rational drug discovery programmes. The
XX antisense nucleic acid sequence is also useful to screen for homologous
XX nucleic acids which are required for cell proliferation in a wide variety
XX of organisms. The present sequence represents an essential prokaryotic
XX cellular proliferation protein. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 151 AA;

Query Match 5.7%; Score 9; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 KPFFPGVLI 70

DB 58 KPFFPGVLI 66

RESULT 24

ABU15025
ID ABU15025 standard; protein; 151 AA.

XX AC ABU15025;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #552.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Escherichia coli.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlseen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI: 2003-029926/02.
XX N-PSDB; ACA18895.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 42949; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
XX the 5213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than S. aureus, S. typhimurium,
XX K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 151 AA;

Query Match 5.7%; Score 9; DB 6; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 KPFFPGVLI 70

DB 58 KPFFPGVLI 66

RESULT 25

ABU23117
 ID ABU23117 standard; protein; 151 AA.
 AC ABU23117;
 DT 19-JUN-2003 (first entry)
 XX Protein encoded by Prokaryotic essential gene #8644.
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Bordetella pertussis.
 OS WO200277183-A2.
 XX PD 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX N-PSDB; ACA26987.
 DR WPI; 2003-029926/02.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 51041; 1766pp; English.
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC on a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 151 AA;

Query Match 5.7%; Score 9; DB 6; Length 151;
 Best Local Similarity 100.0%; Fred. No. 0.53;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 LPHRYPMLL 29
 |||||
 Db 12 LPHRYPMLL 20
 RESULT 26
 ADN18054
 ID ADN18054 standard; protein; 151 AA.
 XX
 AC ADN18054;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #707.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 XX 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINKLE/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 DR WPI; 2004-061375/06.
 XX
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 707; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan

CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 151 AA;

Query Match 5.7%; Score 9; DB 8; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIFPFGVLI 70
Db 58 KPIFPFGVLI 66
|||||

RESULT 27
ADN17554
ID ADN17554 standard; protein; 157 AA.
XX
AC ADN17554;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #207.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
XX Bacteria.
OS
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
DR
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 207; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC

CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 157 AA;

Query Match 5.7%; Score 9; DB 8; Length 157;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIFPFGVLI 70
Db 64 KPIFPFGVLI 72
|||||

RESULT 28
ADF06342
ID ADF06342 standard; protein; 162 AA.
XX
AC ADF06342;
XX
DT 12-FEB-2004 (first entry)
XX
DE Bacterial polypeptide #2455.
XX
KW Proteus mirabilis infection; bacterial infection; antibacterial;
XX immunostimulant.
XX
OS Proteus mirabilis.
XX
PN US6605709-B1.
XX
PD 12-AUG-2003.
XX
PF 05-APR-2000; 2000US-00543681.
XX
PR 09-APR-1999; 99US-0128706P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL;
XX
DR WPI; 2003-895291/82.
DR N-PSDB; ADF02170.
XX
XX New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.
XX
PS Disclosure; SEQ ID NO 6627; 870pp; English.
XX
XX The invention relates to new Proteus mirabilis polypeptides and
CC polynucleotides. The invention also relates to antibodies against the
CC polypeptides, methods for producing the polypeptides, a method of
CC generating vaccines for immunising an individual against P. mirabilis, a
CC method for evaluating a compound for the ability to bind a P. mirabilis
CC polypeptide and a method for screening test compounds for anti-bacterial
CC activity. The polypeptides and polynucleotides are useful as molecular
CC targets for diagnosing, preventing and treating pathological conditions
CC resulting from bacterial infection, as reagents for diagnosis of
CC bacterial diseases, as components of antibacterial vaccines, as targets
CC for antibacterial drugs or as bio-control agents for plants. This
CC sequence represents a Proteus mirabilis polypeptide of the invention.
XX

SQ Sequence 162 AA;

Query Match 5.7%; Score 9; DB 7; Length 162;
Best Local Similarity 100.0%; Pred. No. 0.57;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KPIPPGVLI 70
Db 69 KPIPPGVLI 77

RESULT 29
ABU40649
ID ABU40649 standard; protein; 172 AA.
XX
AC ABU40649;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #26176.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Proteus sp.
XX
FN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
WPI; 2003-029926/02.
DR N-PSDB; ACA44519.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
Claim 25; SEQ ID NO 68573; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC on a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 172 AA;

Query Match 5.7%; Score 9; DB 6; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KPIPPGVLI 70
Db 81 KPIPPGVLI 89

RESULT 30
ADS42775
ID ADS42775 standard; protein; 172 AA.
XX
AC ADS42775;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #21205.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
FN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PF 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
Claim 1; SEQ ID NO 21205; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 172 AA;

Query Match 5.7%; Score 9; DB 8; Length 172;
 Best Local Similarity 100.0%; Pred. No. 0.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFPFGVLI 70

Db 79 KPFPFGVLI 87

RESULT 31

ABUS0043

ID ABUS0043 standard; protein; 181 AA.

XX AC

XX ABUS0043;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #35570.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Yersinia pestis.

XX WO200277183-A2.
 XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.
 XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.
 XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.
 XX N-PSDB; ACA53913.

XX New antisense nucleic acids, useful for identifying proteins or screening
 XX for homologous nucleic acids required for cellular proliferation to
 XX isolate candidate molecules for rational drug discovery programs.
 PS Claim 25; SEQ ID NO 77967; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway; (8)
 CC required for proliferation, or that inhibits cellular proliferation; (9)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC on a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 181 AA;

Query Match 5.7%; Score 9; DB 6; Length 181;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFPFGVLI 70

Db 58 KPFPFGVLI 66

RESULT 32

AAY90363

ID AAY90363 standard; protein; 133 AA.

XX AC

XX AAY90363;

XX 12-SEP-2003 (revised)

XX 15-JAN-2001 (first entry)

XX *S. zoepidemicus* gene clone #2 protein #1.
 XX DNA cloning; gene cloning; gene detection.

XX *Streptococcus equi* subsp. *zoepidemicus*.
 XX EP1035202-A2.

XX 13-SEP-2000.

XX 09-MAR-2000; 2000EP-00104485.

XX 11-MAR-1999; 99JP-00064432.

XX (JCRP-) JCR PHARM CO LTD.

XX Hong K;

XX WPI; 2000-559872/52.

XX N-PSDB; AAA37788.

XX Cloning unknown gene fragments by introducing cloning vector with ligated
 XX DNA PCR products containing restriction enzyme cleaved and
 XX dephosphorylated plasmid DNA and chromosomal DNA into competent cells.
 PS Claim 5; Page 10-12; 14pp; English.

XX This sequence represents a protein encoded by a gene clone from
CC Streptococcus zooepidemicus. This sequence was cloned using the method of
CC the invention. The method is for cloning DNA fragments, and comprises:
CC (a) cleaving a plasmid vector DNA with a restriction enzyme; (b)
CC dephosphorylating an end of the cleaved DNA; (c) obtaining a mixture of
CC restriction enzyme cleaved fragments of chromosomal DNA from a given
CC organism using the same restriction enzyme as for the plasmid vector; (d)
CC obtaining a mixture of ligated DNA's through the ligation of the plasmid
CC DNA and dephosphorylated chromosomal DNA fragments; (e) PCR amplifying
CC the mixture of ligated DNA's using arbitrary primers; and (f) introducing
CC a cloning vector containing the PCR product into a competent cell. The
CC method can be used to clone fragments of an unknown gene using arbitrary
CC primers. Cloning unknown genes using arbitrary primers enables the
CC selection of genes, which allows efficient detection of new genes that
CC may have potential applications such as therapeutics. (Updated on 12-SEP-
CC 2003 to standardise OS field)
XX
SQ Sequence 133 AA;
Query Match 5.0%; Score 8; DB 3; Length 133;
Best Local Similarity 100.0%; Pred. No. 5,6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26 PMLLVDR 33
Db 9 PMLLVDR 16
|||||
RESULT 33
ADN20318
ID ADN20318 standard; protein; 137 AA.
AC ADN20318;
XX
XX 02-DEC-2004 (first entry)
XX Bacterial polypeptide #2971.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
XX US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 2971; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved lignin production or improved galactomannan
CC condition. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 137 AA;
Query Match 5.0%; Score 8; DB 8; Length 137;
Best Local Similarity 100.0%; Pred. No. 5,7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 66 PGVLIVEG 73
Db 54 PGVLIVEG 61
|||||
RESULT 34
ABU30001
ID ABU30001 standard; protein; 139 AA.
AC ABU30001;
XX
XX 19-JUN-2003 (first entry)
XX Protein encoded by prokaryotic essential gene #15528.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Enterococcus faecium.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlseen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX N-PSDB; ACA33871.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 57925; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of

XX CC the 6213 antisense sequences given in the specification where expression

XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:

XX CC (1) a vector comprising a promoter operably linked to the nucleic acid

XX CC encoding a polypeptide whose expression is inhibited by the antisense

XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX CC polypeptide or its fragment whose expression is inhibited by the

XX CC antisense nucleic acid; (4) an antibody capable of specifically binding

XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

XX CC proliferation or the activity of a gene in an operon required for

XX CC proliferation; (7) identifying a compound that influences the activity of

XX CC the gene product or that has an activity against a biological pathway

XX CC required for proliferation, or that inhibits cellular proliferation; (8)

XX CC identifying a gene required for cellular proliferation or the biological

XX CC pathway in which a proliferation-required gene or its gene product lies

XX CC or a gene on which the test compound that inhibits proliferation of an

XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

XX CC compound's activity; (11) a culture comprising strains in which the gene

XX CC product is overexpressed or underexpressed; (12) determining the extent

XX CC to which each of the strains is present in a culture or collection of

XX CC strains; or (13) identifying the target of a compound that inhibits the

XX CC proliferation of an organism. The antisense nucleic acids are useful for

XX CC identifying proteins or screening for homologous nucleic acids required

XX CC for cellular proliferation to isolate candidate molecules for rational

XX CC drug discovery programs, or for screening homologous nucleic acids

XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

XX CC the target prokaryotic essential genes. Note: The sequence data for this

XX CC patent did not form part of the printed specification, but was obtained

XX CC in electronic format directly from WIPO at

XX CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 139 AA;

Query Match 5.0%; Score 8; DB 6; Length 139;

Best Local Similarity 100.0%; Pred. No. 5.8;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 PHRYPMLL 29

DB 11 PHRYPMLL 18

RESULT 35

ABU39542

ID ABU39542 standard; protein; 152 AA.

XX AC ABU39542;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #25069.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Pasteurella multocida.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX XX (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA43412.

XX PT New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 67466; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of

XX CC the 6213 antisense sequences given in the specification where expression

XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:

XX CC (1) a vector comprising a promoter operably linked to the nucleic acid

XX CC encoding a polypeptide whose expression is inhibited by the antisense

XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX CC polypeptide or its fragment whose expression is inhibited by the

XX CC antisense nucleic acid; (4) an antibody capable of specifically binding

XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

XX CC proliferation or the activity of a gene in an operon required for

XX CC proliferation; (7) identifying a compound that influences the activity of

XX CC the gene product or that has an activity against a biological pathway

XX CC required for proliferation, or that inhibits cellular proliferation; (8)

XX CC identifying a gene required for cellular proliferation or the biological

XX CC pathway in which a proliferation-required gene or its gene product lies

XX CC or a gene on which the test compound that inhibits proliferation of an

XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

XX CC compound's activity; (11) a culture comprising strains in which the gene

XX CC product is overexpressed or underexpressed; (12) determining the extent

XX CC to which each of the strains is present in a culture or collection of

XX CC strains; or (13) identifying the target of a compound that inhibits the

XX CC proliferation of an organism. The antisense nucleic acids are useful for

XX CC identifying proteins or screening for homologous nucleic acids required

XX CC for cellular proliferation to isolate candidate molecules for rational

XX CC drug discovery programs, or for screening homologous nucleic acids

XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

XX CC the target prokaryotic essential genes. Note: The sequence data for this

XX CC patent did not form part of the printed specification, but was obtained

XX CC in electronic format directly from WIPO at

XX CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 152 AA;

Query Match 5.0%; Score 8; DB 6; Length 152;

Best Local Similarity 100.0%; Pred. No. 6.3;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 PIFPGVLI 70

DB 62 PIFPGVLI 69

RESULT 36

ADS30417

ID ADS30417 standard; protein; 154 AA.

XX AC ADS30417;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #19450.

XX KW Recombinant DNA construct; transformed plant; improved plant property;

XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

XX KW pathogen tolerance; pest tolerance; plant disease resistance;

XX KW cell cycle pathway modification; plant growth regulator;

XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;

XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

XX KW bacterial polypeptide.

XX Bacteria.
XX US2003233675-A1.
XX 18-DEC-2003.
XX 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX Claim 1; SEQ ID NO 19450; 122pp; English.
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX homologous recombination, modified seed oil or protein yield and/or
XX content, improved yield by modification of carbohydrate, nitrogen or
XX phosphorus use and/or uptake, by modification of photosynthesis or by
XX providing improved plant growth and development under at least one stress
XX condition, improved lignin production or improved galactomannan
XX production. This sequence represents a bacterial polypeptide used in the
XX scope of the invention. Note: The sequence data for this patent did not
XX form part of the printed specification but was obtained in electronic
XX format from USPTO at seqdata.uspto.gov/sequence.html.
XX Sequence 154 AA;
SQ Query Match 5.0%; Score 8; DB 8; Length 154;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 LPHRYPML 28
Db 17 LPHRYPML 24
RESULT 37
AAG12580
ID AAG12580 standard; protein; 161 AA.
XX
XX AAG12580;
XX
XX 17-OCT-2000 (first entry)
XX
XX Zea mays protein fragment SEQ ID NO: 11749.
XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
OS Zea mays subsp. mays.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132407P.
XX 05-MAY-1999; 99US-0132484P.
XX 06-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
OS Zea mays subsp. mays.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132407P.
XX 05-MAY-1999; 99US-0132484P.
XX 06-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144333P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151068P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.

PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 5.0%; Score 8; DB 3; Length 161;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GFDPEIAK 95
|||
Db 97 GFDPEIAK 104

RESULT 38
ADX77098

ID ADX77098 standard; protein; 437 AA.

XX AC ADX77098;

XX 21-APR-2005 (first entry)

DE Plant full length insert polypeptide seqid 46464.

XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;

KW protein content.
OS Unidentified.
PN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
XX 28-APR-2003; 2003US-00425114.
XX
XX 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
XX (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAS/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
DR
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX Claim 1; SEQ ID NO 46464; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
XX Sequence 437 AA;
SQ
Query Match 5.0%; Score 8; DB 8; Length 437;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 88 GFDPEIAK 95
DB 160 GFDPEIAK 167
|||||
RESULT 39
ADN19962
ID ADN19962 standard; protein; 478 AA.
XX
XX ADN19962;
AC
XX
XX 02-DEC-2004 (first entry)
DT
XX Bacterial polypeptide #2615.
DE
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW

KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
XX US2003233675-A1.
PN
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
PF
XX 21-FEB-2002; 2002US-0360039P.
PR
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI
XX WPI; 2004-061375/06.
DR
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 2615; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at segdata.uspto.gov/sequence.html.
XX
XX Sequence 478 AA;
SQ
Query Match 5.0%; Score 8; DB 8; Length 478;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 88 GFDPEIAK 95
DB 219 GFDPEIAK 226
|||||
RESULT 40
ADS43079
ID ADS43079 standard; protein; 910 AA.
XX
XX ADS43079;
AC

Search completed: January 10, 2006, 13:39:47
Job time : 78 secs

XX 02-DEC-2004 (first entry)
XX Bacterial polypeptide #21509.
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX Bacteria.
OS Bacteria.
XX US2003233675-A1.
XX 18-DEC-2003.
XX 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX Claim 1; SEQ ID NO 21509; 122pp; English.
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 910 AA;
SQ Query Match 5.0%; Score 8; DB 8; Length 910;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 147 AEAELKAM 154
Db 424 AEAELKAM 431

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:36:51 ; Search time 16 Seconds
(without alignments)
956.154 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 159

Sequence: 1 MEQSHQNLSQFFIEHILQI.....QVDGKVAEAEELKAMIAERD 159

Scoring table: OLIGO 3
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 8

Total number of hits satisfying chosen parameters: 18

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : PIR 80:*

1: piri:*

2: piri2:*

3: piri3:*

4: piri4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	159	100.0	159	2 H64691	(3R)-hydroxymyrist
2	118	74.2	159	2 H71826	(3R)-hydroxymyrist
3	12	7.5	140	2 B95049	hypothetical prote
4	12	7.5	140	2 H97919	(3R)-hydroxymyrist
5	11	6.9	141	2 C97338	hydroxymyristoyl-(
6	10	6.3	146	2 H81445	(3R)-hydroxymyrist
7	9	5.7	151	2 D64742	(3R)-hydroxymyrist
8	9	5.7	151	2 F85502	(3R)-hydroxymyrist
9	5.7	151	2 F90651		hypothetical prote
10	9	5.7	181	2 AF0129	(3R)-hydroxymyrist
11	9	5.7	201	2 S35968	(3R)-hydroxymyrist
12	8	5.0	105	2 S55691	keratin type I (cl
13	8	5.0	137	2 B72335	(3R)-hydroxymyrist
14	8	5.0	368	2 A28825	keratin, type I no
15	8	5.0	478	2 S74603	dihydrolipoamide d
16	8	5.0	613	2 G69429	Glu-tRNA amidotran
17	8	5.0	841	2 T01011	hypothetical prote
18	8	5.0	910	2 C69069	cation-transportin

ALIGNMENTS

RESULT 1

H64691

(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Helicobacter pyl

C;Species: Helicobacter pylori

C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004

C;Accession: H64691

R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997

A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A;Reference number: A64520; MUID:97394467; PMID:9252185

A;Accession: H64691

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-159 <TOM>

A;Cross-references: UNIPROT:O25928; UNIPARC:UPI000012A4AD; GB:AE000637; GB:AE000511; NID:

C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 100.0%; Score 159; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.2e-158;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEQSHQNLSQFFIEHILQIILPHRYPMLLVDRRIIELOANKKIVAYKNITFNEDVFNHGHP 60

Db 1 MEQSHQNLSQFFIEHILQIILPHRYPMLLVDRRIIELOANKKIVAYKNITFNEDVFNHGHP 60

Qy 61 NKPIPPGVLIVEGMAQTGGFLAFTSLMGDFDEIAKTKIVYFMTIDKVKRIPVTPGDRLE 120

Db 61 NKPIPPGVLIVEGMAQTGGFLAFTSLMGDFDEIAKTKIVYFMTIDKVKRIPVTPGDRLE 120

Qy 121 YHLEVLKHKGMIMQVGGTAQVDGKVAEAEELKAMIAERD 159

Db 121 YHLEVLKHKGMIMQVGGTAQVDGKVAEAEELKAMIAERD 159

RESULT 2

H71826

(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Helicobacter pyl

C;Species: Helicobacter pylori

A;Variety: strain J99

C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004

C;Accession: H71826

R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; J

Nature 397, 176-180, 1999

A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A;Reference number: A71800; MUID:99120557; PMID:9923682

A;Accession: H71826

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-159 <ARN>

A;Cross-references: UNIPROT:Q9ZJL6; UNIPARC:UPI000012A4AC; GB:AE001551; GB:AE001439; NID:

A;Experimental source: strain J99

C;Genetics:

A;Gene: fabZ

C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

C;Keywords: carbon-oxygen lyase; hydro-lyase; hydrolase

Query Match 74.2%; Score 118; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 9.9e-116;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 KIVAYKNITFNEDVFNHGHPFNKPIPPGVLIVEGMAQTGGFLAFTSLMGDFDPEIAKTKIVY 100

Db 41 KIVAYKNITFNEDVFNHGHPFNKPIPPGVLIVEGMAQTGGFLAFTSLMGDFDPEIAKTKIVY 100

Qy 101 FMTIDKVKFRIPVTPGDRLEHVLKHKGMIMQVGGTAQVDGKVAEAEELKAMIAER 158

Db 101 FMTIDKVKFRIPVTPGDRLEHVLKHKGMIMQVGGTAQVDGKVAEAEELKAMIAER 158

RESULT 3

B95049

hypothetical protein SP0424 [imported] - Streptococcus pneumoniae (strain TIGR4)

C;Species: Streptococcus pneumoniae

C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004

C;Accession: B95049
R;Tetrelan, H.; Neilson, K.E.; Paulsen, I.T.; Eissen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: B95049
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-140 <KUR>
A;Cross-references: UNIPROT:P59201; UNIPARC:UPI0000051423; GB:AE005672; PIDN:AAK74587.1;
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0424
C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

Query Match 7.5%; Score 12; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 8.6e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
|||||
Db 11 LPHRYPMLLVDR 22

RESULT 4
H97919
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) [imported] - Stre
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: H97919
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; B
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: H97919
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-140 <KUR>
A;Cross-references: UNIPROT:P59202; UNIPARC:UPI000012A4BA; GB:AE007317; PIDN:AAK99188.1;
C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
C;Keywords: carbon-oxygen lyase; hydro-lyase
C;Genetics:
A;Gene: fabZ

Query Match 7.5%; Score 12; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 8.6e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
|||||
Db 11 LPHRYPMLLVDR 22

RESULT 5
C97338
hydroxymyristoyl-(acyl carrier protein) dehydratase [imported] - Clostridium acetobutyli
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: C97338
R;Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: C97338
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-141 <KUR>
A;Cross-references: UNIPROT:Q97DA9; UNIPARC:UPI0000012A4A6; GB:AE001437; PIDN:AAK81494.1;
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3571
C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

Query Match 6.9%; Score 11; DB 2; Length 141;
Best Local Similarity 100.0%; Pred. No. 0.00097;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 PHRYPMLLVDR 32
|||||
Db 13 PHRYPMLLVDR 23

RESULT 6
H81445
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) Cj0273 [imported] -
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: H81445
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Bagham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A;Reference number: A81250; MUID:20150912; PMID:10688204
A;Accession: H81445
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-146 <PAR>
A;Cross-references: UNIPROT:Q9PIM2; UNIPARC:UPI0000012A4A2; GB:AL139074; GB:AL111168; NID:
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: fabZ; Cj0273
C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 6.3%; Score 10; DB 2; Length 146;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VAEAEKAKMI 155
|||||
Db 134 VAEAEKAKMI 143

RESULT 7
D64742
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Escherichia coli
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: D64742; B33171
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: D64742
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-151 <BLAT>
A;Cross-references: UNIPROT:P21774; UNIPARC:UPI0000012A4A9; GB:AE000127; GB:U00096; NID:G
A;Experimental source: strain K-12, substrain MGL655
R;Coleman, J.; Raetz, C.R.H.
J. Bacteriol. 170, 1268-1274, 1988
A;Title: First committed step of lipid A biosynthesis in Escherichia coli: sequence of th
A;Reference number: A33171; MUID:88139188; PMID:3277952
A;Accession: B33171
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-139, 'DR', 142-149, 'RRDT', <COL>
A;Cross-references: UNIPARC:UPI0000178D50; GB:M19334; GB:M18265; GB:M18266

```

C;Genetics:
A;Gene: fabZ
A;Start codon: TTG
C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match      5.7%; Score 9; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFPFGVLI 70
Db 58 KPFPFGVLI 66

RESULT 8
F85502
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) [similarity] - Esc
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: F85502
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamouis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F85502
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-151 <STO>
A;Cross-references: UNIPROT:P21774; UNIPARC:UPI000012A4A9; GB:AB005174; NID:gl2512908; E
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: fabZ
C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match      5.7%; Score 9; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFPFGVLI 70
Db 58 KPFPFGVLI 66

RESULT 9
F90651
hypothetical protein ECs0182 [imported] - Escherichia coli (strain O157:H7, substrain R)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F90651
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90651
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-151 <HAY>
A;Cross-references: UNIPROT:P21774; UNIPARC:UPI00001653A3; GB:BA000007; PIDN:BA033605.1;
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECs0182
C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase

Query Match      5.7%; Score 9; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFPFGVLI 70
Db 58 KPFPFGVLI 66

RESULT 10
AF0129
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) [impor
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AF0129
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; E
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, E
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AF0129
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-181 <KUR>
A;Cross-references: UNIPROT:Q8ZH57; UNIPARC:UPI000012A4C3; GB:AL590842; PIDN:CAC89897.1;
C;Genetics:
A;Gene: fabZ
C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match      5.7%; Score 9; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFPFGVLI 70
Db 58 KPFPFGVLI 66

RESULT 11
S35968
(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Yersinia enterocoli
C;Species: Yersinia enterocolitica
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S35968
R;Vuorio, R.; Harkonen, T.; Vaara, M.
submitted to the EMBL Data Library, August 1993
A;Reference number: S35968
A;Accession: S35968
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-201 <VUU>
A;Cross-references: UNIPROT:P32205; UNIPARC:UPI00001705F3; EMBL:Z25463; NID:g397536; PIDN
C;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match      5.7%; Score 9; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFPFGVLI 70
Db 81 KPFPFGVLI 89

RESULT 12
S55691
keratin type I (clone Jn7-4) - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 28-Oct-1995 #sequence_revision 07-Feb-1997 #text_change 09-Jul-2004
C;Accession: S55691
R;Adati, N.; Ito, T.; Koga, C.; Kito, K.; Sakaki, Y.; Shiokawa, K.
Biochim. Biophys. Acta 1262, 43-51, 1995
A;Title: Differential display analysis of gene expression in developing embryos of Xenopi
A;Reference number: S55691; MUID:95290493; PMID:7772598
A;Accession: S55691
A;Molecule type: mRNA
A;Residues: 1-105 <ADA>

```

A;Cross-references: UNIPROT:Q91404; UNIPARC:UPI00000FC2C7; EMBL:S78089; NID:g999222; PID: C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil; intermediate filament

Query Match 5.0%; Score 8; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 VDGVKVAE 148
Db 89 VDGVKVAE 96
|||||

RESULT 13
B72335
A;Title: Hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-) - Thermotoga maritima
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: B72335
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: B72335
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-137 <R>
A;Cross-references: UNIPROT:Q9WZQ8; UNIPARC:UPI000012A4BC; GB:AE001747; GB:AE000512; NID: C;Experimental source: strain MSB8
C;Genetics:
A;Gene: TW0801
A;Superfamily: (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 5.0%; Score 8; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 PGVLIVEG 73
Db 54 PGVLIVEG 61
|||||

RESULT 14
A28825
keratin, type I nonepidermal - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C;Accession: A28825
R;LaFlamme, S.E.; Jamrich, M.; Richter, K.; Sargent, T.D.; Dawid, I.B.
Genes Dev. 2, 853-862, 1988
A;Title: Xenopus endo B is a keratin preferentially expressed in the embryonic notochord
A;Reference number: A28825; MUID:89092007; PMID:2463213
A;Accession: A28825
A;Molecule type: mRNA
A;Residues: 1-368 <LAP>
A;Cross-references: UNIPROT:P08802; UNIPARC:UPI000012DAEC; GB:Y00230; NID:g64863; PIDN: C;Genetics:
A;Start codon: GGT
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil

Query Match 5.0%; Score 8; DB 2; Length 368;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 VDGVKVAE 148
Db 352 VDGVKVAE 359
|||||

RESULT 15
S74603
dihydrolipoamide dehydrogenase phdD - Synechocystis sp. (strain PCC 6803)
N;Alternate names: hypothetical protein slri096
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 03-Nov-2000
C;Accession: S74603
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp.
S.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74603
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-478 <KAN>
A;Cross-references: UNIPARC:UPI0000164C2B; EMBL:D90900; GB:AB001339; NID:g1651768; PIDN: F;14-470/Domain: dihydrolipoamide dehydrogenase homology <DLD>
F;48-53/Disulfide bonds: redox-active #status predicted
C;Genetics:
A;Gene: phdD
A;Start codon: GTG
C;Superfamily: dihydrolipoamide dehydrogenase; dihydrolipoamide dehydrogenase homology
C;Keywords: redox-active disulfide
F;14-470/Domain: dihydrolipoamide dehydrogenase homology <DLD>
F;48-53/Disulfide bonds: redox-active #status predicted

Query Match 5.0%; Score 8; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GFDPEIAK 95
Db 219 GFDPEIAK 226
|||||

RESULT 16
G69429
Glu-tRNA amidotransferase, subunit B (gatB-1) homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: G69429
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.A. Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: G69429
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-613 <LE>
A;Cross-references: UNIPROT:O28832; UNIPARC:UPI0000056D1C; GB:AE001004; GB:AE000782; NID: C;Genetics:
A;Start codon: GGT
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil

Query Match 5.0%; Score 8; DB 2; Length 613;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 VDGVKVAE 148
Db 594 VDGVKVAE 601
|||||

RESULT 17
T01011
hypothetical protein At2g39810 [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein T517.11
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01011; G84821

R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, November 1997
A:Description: Arabidopsis thaliana chromosome II BAC T517 genomic sequence.
A:Reference number: Z14162
A:Accession: T01011
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-841 <ROU>
A:Cross-references: UNIPROT:O22289; UNIPARC:UPI000000AB839; EMBL:AC003000; NID:g2642152;
A:Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84821
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-841 <STO>
A:Cross-references: UNIPARC:UPI000000AB839; GB:AE002093; NID:g2642163; PIDN:AAB87130.1; G
C:Genetics:
A:Gene: T517.11; At2g39810
A:Map position: 2
A:Introns: 133/2; 200/3; 238/3; 365/3; 529/3; 575/3; 788/2
C:Superfamily: Arabidopsis thaliana hypothetical protein At2g39810

Query Match 5.0%; Score 8; DB 2; Length 841;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 LVDRIIEL 36
Db 486 LVDRIIEL 493

RESULT 18
C69069
cation-transporting P-ATPase PacL - Methanobacterium thermoautotrophicum (strain Delta H
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 31-Dec-2004
C:Accession: C69069
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
Kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: C69069
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-910 <MTH>
A:Cross-references: UNIPROT:O27560; UNIPARC:UPI0000066763; GB:AE000912; GB:AE000666; NID
C:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1516
A:Start codon: GTG
C:Superfamily: Na(+)/K(+)-transporting ATPase alpha chain; ATPase nucleotide-binding dom
F:533-704/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 5.0%; Score 8; DB 2; Length 910;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 147 AEAELKAM 154
Db 424 AEAELKAM 431

Search completed: January 10, 2006, 13:41:35
Job time : 16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:35:56 ; Search time 71 Seconds

(without alignments)
1579.986 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 159

Sequence: 1 MEQSHQLSQPFIEHLQI.....QVDGKVVAEELKAMIAERD 159

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2166443 seqs, 705528306 residues

Word size : 8

Total number of hits satisfying chosen parameters: 70

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Uniprot 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	159	100.0	159	1	FABZ HELPY	O25928 helicobacte
2	118	74.2	159	1	FABZ HELPY	O9zj16 helicobacte
3	81	50.9	159	2	O5G940 HELPY	O5g940 helicobacte
4	14	8.8	169	1	FABZ HELPY	O7u319 helicobacte
5	13	8.2	133	2	O9FAV8_9STPE	O9fav8 streptococc
6	13	8.2	142	1	FABZ CLOTE	O899n7 clostridium
7	12	7.5	139	1	FABZ STRP3	O8k631 streptococc
8	12	7.5	139	1	FABZ STRP6	O5xae7 streptococc
9	12	7.5	139	1	FABZ STRP8	P64111 streptococc
10	12	7.5	139	1	FABZ STRPY	P64110 streptococc
11	12	7.5	140	1	FABZ STRA3	O8e720 streptococc
12	12	7.5	140	1	FABZ STRA5	O8elk3 streptococc
13	12	7.5	140	1	FABZ STRPN	P59201 streptococc
14	12	7.5	140	1	FABZ STRR6	P59202 streptococc
15	12	7.5	141	1	FABZ STRT1	O5m177 streptococc
16	12	7.5	141	1	FABZ STRT2	O5m5r4 streptococc
17	12	7.5	164	1	FABZ RALSO	O8xx10 ralstonia s
18	11	6.9	116	2	O7PA53 FUSNV	O7p453 fusobacteri
19	11	6.9	141	1	FABZ CLOAB	O97da9 clostridium
20	10	6.3	142	2	O4HG66 CAMCO	O4hg66 campylobact
21	10	6.3	142	2	O4HJ25 CAMLA	O4hj25 campylobact
22	10	6.3	146	1	FABZ CANJE	O9pkm2 campylobact
23	10	6.3	146	1	FABZ CANJR	O5hwj3 campylobact
24	10	6.3	146	2	O4HOA3 CAMUP	O4hga3 campylobact
25	9	5.7	141	1	FABZ THETN	O8rd71 thermoanaer
26	9	5.7	150	1	FABZ PHOLL	O7n8n6 photorhabdu
27	9	5.7	151	1	FABZ BORBR	O7wj83 bordetella
28	9	5.7	151	1	FABZ BORPA	O7wa49 bordetella
29	9	5.7	151	1	FABZ BORPE	O7wv59 bordetella
30	9	5.7	151	1	FABZ CANBF	O7vrds candidatus
31	9	5.7	151	1	FABZ ECO57	P0a6q8 escherichia

32	9	5.7	151	1	FABZ_ECOL6	P0a6q7 escherichia
33	9	5.7	151	1	FABZ_ECOLI	P0a6q6 escherichia
34	9	5.7	151	1	FABZ_ERWCT	O6d8d2 erwinia car
35	9	5.7	151	1	FABZ_SHIFL	P0a6q9 shigella fl
36	9	5.7	151	1	FABZ_WOLSU	O7mas2 wolinnella s
37	9	5.7	176	1	FABZ_YERPS	O667k0 versinia ps
38	9	5.7	178	1	FABZ_YEREN	P32205 versinia en
39	9	5.7	181	1	FABZ_YERPE	O8zh57 versinia pe
40	9	5.7	298	2	O8YOY7_RALSO	O8yo77 ralstonia s
41	8	5.0	51	2	O4YGS7_PLABE	O4ygs7 plasmodium
42	8	5.0	94	2	O90X75_GILMI	O90x75 gillichthys
43	8	5.0	105	2	O91404_XENLA	O91404 xenopus lae
44	8	5.0	137	1	FABZ_THENA	O9wzq8 thermotoga
45	8	5.0	142	1	FABZ_SYMTH	O6tf91 symbiobacte
46	8	5.0	150	2	O5E3F0_VIBF1	O5e3f0 vibrio fisc
47	8	5.0	151	1	FABZ_IDILO	O5f0z2 idiomarina
48	8	5.0	152	1	FABZ_PASMU	O9cjk9 pasteurella
49	8	5.0	154	1	FABZ_RHIME	O92q46 rhizobium m
50	8	5.0	159	2	O9T1E3_9CAUD	O9tfk3 lactobacill
51	8	5.0	229	2	O745H9_MYCPA	O745h9 mycobacteri
52	8	5.0	326	2	O6M145_METMP	O6m145 methanococc
53	8	5.0	368	1	K1C18_XENLA	P08802 xenopus lae
54	8	5.0	394	2	O518P3_ENTHI	O518p3 entamoeba h
55	8	5.0	428	2	O61NH6_XENLA	O6inh6 xenopus lae
56	8	5.0	432	2	O7SY65_XENLA	O7sy65 xenopus lae
57	8	5.0	436	2	O8AV12_XENLA	O8av12 xenopus lae
58	8	5.0	473	1	DLDH_SYNY3	P72740 synechocyst
59	8	5.0	561	2	O94CN9_ORYSA	O94cn9 oryza sativ
60	8	5.0	590	2	O75M05_ORYSA	O75m05 oryza sativ
61	8	5.0	595	2	O8JTV1_LSDV	O8jtv1 lumpy skin
62	8	5.0	596	2	O91MW7_LSDV	O91mw7 lumpy skin
63	8	5.0	613	1	GATE_ARCFU	O28832 archaeoglob
64	8	5.0	642	2	O50VK5_ENTHI	O50vk5 entamoeba h
65	8	5.0	684	2	O82J56_STRAW	O82j56 streptomyce
66	8	5.0	841	2	O22289_ARATH	O22289 arabidopsis
67	8	5.0	856	2	O4HMA4_CAMCO	O4hbm4 campylobact
68	8	5.0	910	2	O27560_METTH	O27560 methanobact
69	8	5.0	927	2	O84JU6_ARATH	O84ju6 arabidopsis
70	8	5.0	954	2	O54LX3_DICDI	O54lx3 dictyosteli

ALIGNMENTS

RESULT 1

FABZ_HELPI	STANDARD;	PRT;	159 AA.
ID_FABZ_HELPI			
AC_O25928;			
DT_16-OCT-2001 (Rel. 40, Created)			
DT_16-OCT-2001 (Rel. 40, Last sequence update)			
DT_10-MAY-2005 (Rel. 47, Last annotation update)			
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))			
DE ((3R)-hydroxymyristoyl ACP dehydratase).			
GN Name=fabz; OrderedLocusNames=HPI1376;			
OS Helicobacter pylori (Campylobacter pylori).			
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;			
OC Helicobacteraceae; Helicobacter.			
OX NCBI_TaxID=210;			
RN [1]			
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC STRAIN=26695 / ATCC 700392;			
RA MEDLINE=97394467; PubMed=9252185; DOI=10.1038/41483;			
RA Tomb J.-P., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,			
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S.R.,			
RA Dougherty B.A., Nelson K.E., Quackenbush J., Zhou L., Kirkness E.F.,			
RA Peterson S.N., Loftus B.J., Richardson D.L., Dodson R.J., Khalak H.G.,			
RA Glodek A., McKenney K., Fitzgerald L.M., Lee N., Adams M.D.,			
RA Hickey E.K., Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D.,			
RA Kelley J.M., Cotton M.D., Weidman J.F., Fujii C., Bowman C.,			
RA Wathey L., Wallin E., Hayes W.S., Borodovsky M., Karp P.D.,			
RA Smith H.O., Fraser C.M., Venter J.C.;			
RT "The complete genome sequence of the gastric pathogen Helicobacter			
RT pylori."			

RL Nature 388:539-547(1997).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC -----
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AB000637; AAD08419.1; -; Genomic_DNA.
DR FIR; H64691; H64691.
DR HAMAP; MF 00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 58 By similarity.
SQ SEQUENCE 159 AA; 18196 MW; 715B6FBD872AE32 CRC64;

Query Match 100.0%; Score 159; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 9.9e-152; Indels 0; Gaps 0;
Matches 159; Conservative 0; Mismatches 0;

QY 1 MEQSHQNLQSQFFIEHILQILPHRYPMLLVDRIELQANKKIYVAYKNITFNEDVFNHGHP 60
DB 1 MEQSHQNLQSQFFIEHILQILPHRYPMLLVDRIELQANKKIYVAYKNITFNEDVFNHGHP 60

QY 61 NKPIPGVLIVGMAQTGGFLAFTSLWGFDPPEIAKTIVFMTIDKVKFRIPVTPGDRLE 120
DB 61 NKPIPGVLIVGMAQTGGFLAFTSLWGFDPPEIAKTIVFMTIDKVKFRIPVTPGDRLE 120

QY 121 YHLEVLKHGMIWQVGTQAQVDGKVAEAEELKAMIAERD 159
DB 121 YHLEVLKHGMIWQVGTQAQVDGKVAEAEELKAMIAERD 159

RESULT 2
FABZ_HELPJ STANDARD; PRT; 159 AA.
AC Q9ZJL6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-))
DE Names=fabZ; OrderedLocustNames=JHP1290;
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE-99120557; PubMed-9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Malls S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC -----
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----

DR EMBL; AB001551; AAD06864.1; -; Genomic_DNA.
DR FIR; H71826; H71826.
DR HAMAP; MF 00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 58 By similarity.
SQ SEQUENCE 159 AA; 18210 MW; C0F16FBD868AE39 CRC64;

Query Match 74.3%; Score 118; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 2.2e-110; Indels 0; Gaps 0;
Matches 118; Conservative 0; Mismatches 0;

QY 41 KIVAYKNITFNEDVFNHGHPFNKPIPGVLIVGMAQTGGFLAFTSLWGFDPPEIAKTIVY 100
DB 41 KIVAYKNITFNEDVFNHGHPFNKPIPGVLIVGMAQTGGFLAFTSLWGFDPPEIAKTIVY 100

QY 101 FMTIDKVKFRIPVTPGDRLEHVLKHGMIWQVGTQAQVDGKVAEAEELKAMIAER 158
DB 101 FMTIDKVKFRIPVTPGDRLEHVLKHGMIWQVGTQAQVDGKVAEAEELKAMIAER 158

RESULT 3
QSG940_HELPY PRELIMINARY; PRT; 159 AA.
ID QSG940_HELPY
AC QSG940;
DT 10-MAY-2005 (Tremblrel. 30, Created)
DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)
DE ((3R)-hydroxymyristoyl-acyl carrier protein dehydratase.
GN Names=fabZ;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15967411;
RX STRAIN=SS1;
RA Liu W., Luo C., Han C., Peng S., Yang Y., Yue J., Shen X., Jiang H.;
RT "A new beta-hydroxyacyl-acyl carrier protein dehydratase (FabZ) from
RT Helicobacter pylori: Molecular cloning, enzymatic characterization,
RT and structural modeling."
RL Biochem. Biophys. Res. Commun. 333:1078-1086(2005).
DR EMBL; AV725427; AAM22049.1; -; Genomic_DNA.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0016836; Fatty acid biosynthesis; IEA.
DR GO; GO:0006633; Fatty acid biosynthesis; IEA.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
SQ SEQUENCE 159 AA; 18184 MW; B7F16AAA78315B74 CRC64;

Query Match 50.9%; Score 81; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 4.5e-73;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 GGFLAFTSLWGFDPPEIAKTIVFMTIDKVKFRIPVTPGDRLEHVLKHGMIWQVGG 137
DB 78 GGFLAFTSLWGFDPPEIAKTIVFMTIDKVKFRIPVTPGDRLEHVLKHGMIWQVGG 137

QY 138 TAOVDGKVAEAEELKAMIAER 158
DB 138 TAOVDGKVAEAEELKAMIAER 158

RESULT 4
FABZ_HELPJ STANDARD; PRT; 169 AA.
ID FABZ_HELPJ
AC Q7U3L9;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-))

DE ((3R)-hydroxymyristoyl ACP dehydratase).

GN Name-fabz; OrderedLocNames=HH1181;

OS Helicobacter hepaticus.

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

OC Helicobacteraceae; Helicobacter.

OX NCBI_TaxID=32025;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=ATCC 51449 / 3B1;

RX MEDLINE=22709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;

RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,

RA Bell M., Droege M., Farmann B., Fischer H.-P., Ge Z., Hoerster A.,

RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,

RA Schauer D.B., Shen Z., Weber J., Froesch M., Fox J.G.;

RT "The complete genome sequence of the carcinogenic bacterium Helicobacter hepaticus.";

RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).

CC -1- FUNCTION: Involved in saturated fatty acids biosynthesis.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: Belongs to the thioester dehydratase family.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC -----

DR EMBL; AE01147; AAP77778.1; -; Genomic_DNA.

DR HAMAP; MF 00406; -; 1.

DR InterPro; IPR010084; FabZ.

DR TIGRFAMs; TIGR01750; fabZ; 1.

DR Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.

KW ACT SITE 66 By similarity.

FT ACT SITE 66

SQ SEQUENCE 169 AA; 19071 MW; 011CAB723E36E25C CRC64;

Query Match 8.8%; Score 14; DB 1; Length 169;

Best Local Similarity 100.0%; Pred. No. 1.7e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 QILPHRYPMMLVDR 32

DB 20 QILPHRYPMMLVDR 33

RESULT 5

Q9FAY8_9STRE PRELIMINARY; PRT; 133 AA.

AC Q9FAY8;

DT 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

DE Similar to Bacillus subtilis hydroxymyristoyl-(acyl carrier protein) dehydratase (Fragment).

DE Streptococcus equi subsp. zooepidemicus.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=40041;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=VTU211;

RA Hong K.;

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB030883; BAB16042.1; -; Genomic_DNA.

DR HSSP; P18391; IMKA.

DR GO; GO:0005737; Cytoplasm; IEA.

DR GO; GO:0016836; F:hydro-lyase activity; IEA.

DR GO; GO:0016829; P:lyase activity; IEA.

DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.

DR GO; GO:0009245; P:lipid A biosynthesis; IEA.

DR InterPro; IPR010084; FabZ.

DR TIGRFAMs; TIGR01750; fabZ; 1.

FT NON_TER 1

SQ SEQUENCE 133 AA; 14665 MW; 70PBF56236598A07 CRC64;

Query Match 8.2%; Score 13; DB 2; Length 133;

Best Local Similarity 100.0%; Pred. No. 0.00015;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 LPHRYPMMLLVDR 33

DB 4 LPHRYPMMLLVDR 16

RESULT 6

FABZ_CLOTE STANDARD; PRT; 142 AA.

ID -FABZ_CLOTE STANDARD; PRT; 142 AA.

AC Q899N7;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)

DE ((3R)-hydroxymyristoyl ACP dehydratase).

GN Name=fabZ; OrderedLocNames=CTC00133;

OS Clostridium tetani.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI_TaxID=1513;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=Massachusetts / E88;

RX MEDLINE=22457253; PubMed=12552129; DOI=10.1073/pnas.0335853100;

RA Brueggemann H., Baeumer S., Fricke W.F., Wierzer A., Liesegang H.,

RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,

RA Gottschalk G.;

RT "The genome sequence of Clostridium tetani, the causative agent of tetanus disease.";

RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).

CC -1- FUNCTION: Involved in saturated fatty acids biosynthesis.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: Belongs to the thioester dehydratase family.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.

CC -----

DR EMBL; AE015936; AAC34785.1; -; Genomic_DNA.

DR HAMAP; MF 00406; -; 1

DR InterPro; IPR010084; FabZ.

DR TIGRFAMs; TIGR01750; fabZ; 1.

KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.

FT ACT SITE 50 By similarity.

SQ SEQUENCE 142 AA; 15706 MW; B8D7A9F53E3CF336 CRC64;

Query Match 8.2%; Score 13; DB 1; Length 142;

Best Local Similarity 100.0%; Pred. No. 0.00016;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 ILQILPHRYPMML 29

DB 9 ILQILPHRYPMML 21

RESULT 7

FABZ_STRP3 STANDARD; PRT; 139 AA.

ID -FABZ_STRP3 STANDARD; PRT; 139 AA.

AC Q8K631;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)

DE ((3R)-hydroxymyristoyl ACP dehydratase).

GN Name=fabZ; OrderedLocNames=SPYM3_1520, SP80346;

OS Streptococcus pyogenes (serotype M3).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

Voyich J.M., DeLeo F.R., Martin J.M., Somerville G.A., Musser J.M.;
 "Progress toward characterization of the group A Streptococcus
 metagenome: complete genome sequence of a macrolide-resistant serotype
 M6 strain.";
 J. Infect. Dis. 190:727-738(2004).
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.

EMBL; CP000003; AAAT87616.1; ALT_INIT; Genomic_DNA.
 DR HAMAP; MF 00406; -; 1.
 DR InterPro; IPR010084; FabZ.
 DR TIGRFAMs; TIGR01750; fabz; 1.
 KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
 FT ACT SITE 46 By similarity.
 SQ SEQUENCE 139 AA; 15324 MW; 06090333FC20B479 CRC64;

Query Match 7.5%; Score 12; DB 1; Length 139;
 Best Local Similarity 100.0%; Pred.No. 0.0016;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMMLVDR 32
 |||||
 DB 10 LPHRYPMMLVDR 21

RESULT 9
 FABZ_STRP8 STANDARD; PRT; 139 AA.
 AC P64111; P58175;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
 GN Names=fabZ; OrderedLocNames=spyl1818;
 OS Streptococcus pyogenes (serotype M18);
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=301451;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=MGA8232 / Serotype M18;
 RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
 RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
 RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
 RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
 RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
 RT "Genome sequence and comparative microarray analysis of serotype M18
 group A Streptococcus strains associated with acute rheumatic fever
 outbreaks.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.

EMBL; AE010088; AAL98338.1; -; Genomic_DNA.
 DR HSPSP; P18391; IMKA.
 DR HAMAP; MF 00406; -; 1.
 DR InterPro; IPR010084; FabZ.
 DR TIGRFAMs; TIGR01750; fabz; 1.

KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
 FT ACT_SITE 46 By similarity.
 SQ SEQUENCE 139 AA; 15324 MW; 06090333FC20B479 CRC64;

Query Match 7.5%; Score 12; DB 1; Length 139;
 Best Local Similarity 100.0%; Pred. No. 0.0016;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
 |||||
 DB. 10 LPHRYPMLLVDR 21

RESULT 10

FABZ_STRPY STANDARD; PRT; 139 AA.
 AC P64110; P58175;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
 DE (3R)-hydroxymyristoyl ACP dehydratase).
 GN Name=fabZ; OrderedLocNames=SPY1746;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CC NCBI_TaxID=1314;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
 RX MEDLINE=21192684; PubMed=11296296; DOI=10.1073/pnas.071559398;
 RA Ferrerini J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
 RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the thioester dehydratase family.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; AE006603; AAK34490.1; -; Genomic_DNA.
 DR HSSP; P18391; IMKA.
 DR HAMAP; MF_00406; -; 1.
 DR InterPro; IPR010084; FabZ.
 DR TIGRFAMs; TIGR01750; fabZ; 1.

KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
 FT ACT_SITE 46 By similarity.
 SQ SEQUENCE 139 AA; 15324 MW; 06090333FC20B479 CRC64;

Query Match 7.5%; Score 12; DB 1; Length 139;
 Best Local Similarity 100.0%; Pred. No. 0.0016;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
 |||||
 DB. 10 LPHRYPMLLVDR 21

RESULT 11

FABZ_STR33 STANDARD; PRT; 140 AA.
 AC Q8E720;
 DT 05-JUL-2004 (Rel. 44, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)

DE (3R)-hydroxymyristoyl ACP dehydratase).
 GN Name=fabZ; OrderedLocNames=gbs0338;
 OS Streptococcus agalactiae (serotype III).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CC NCBI_TaxID=216495;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=NEW316 / Serotype III;
 RX MEDLINE=22242508; PubMed=12354221;
 RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Prangeul L.,
 RA Meadek T., Zouline M., Couve E., Lalloui L., Poyart C., Trieu-Cuot P.,
 RA Kunst F.;
 RT "Genome sequence of Streptococcus agalactiae, a pathogen causing invasive neonatal disease";
 RL Mol. Microbiol. 45:1499-1513(2002).
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the thioester dehydratase family.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; AL766844; CAD45983.1; -; Genomic_DNA.
 DR HSSP; P18391; IMKA.
 DR Sagalinst; gbs0338; -.
 DR HAMAP; MF_00406; -; 1.
 DR InterPro; IPR010084; FabZ.
 DR TIGRFAMs; TIGR01750; fabZ; 1.
 KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
 FT ACT_SITE 47 By similarity.
 SQ SEQUENCE 140 AA; 15445 MW; FF12DAED40152268 CRC64;

Query Match 7.5%; Score 12; DB 1; Length 140;
 Best Local Similarity 100.0%; Pred. No. 0.0016;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
 |||||
 DB. 11 LPHRYPMLLVDR 22

RESULT 12

FABZ_STR45 STANDARD; PRT; 140 AA.
 AC Q8E1K3;
 DT 05-JUL-2004 (Rel. 44, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
 DE (3R)-hydroxymyristoyl ACP dehydratase).
 GN Name=fabZ; OrderedLocNames=SAG0351;
 OS Streptococcus agalactiae (serotype V).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CC NCBI_TaxID=216466;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=2603 V/R / ATCC BAA-611 / Serotype V;
 RX MEDLINE=22222988; PubMed=12200547; DOI=10.1073/pnas.182380799;
 RA Tettelin H., Maignani V., Cieslewicz M.J., Eisen J.A., Peterson S.N.,
 RA Wessels L.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
 RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
 RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
 RA Radune D., Fedorova N.B., Scanlan D., Khouri H.M., Mulligan S.,
 RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
 RA Iacobini E.T., Brettoni R., Galli G., Mariani M., Vegni F., Malone D.,
 RA Rinaldo D., Rappuoli R., Telford J.B., Kasper D.L., Grandi G.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative genomic analysis of an

RT emerging human pathogen, serotype V *Streptococcus agalactiae*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL; AE014207; AAM99257.1; -; Genomic_DNA.
 DR HSSP; P18391; IMKA.
 DR TIGR; SAG0351; -.
 DR HAMAP; MF 00406; -; 1.
 DR InterPro; IPRO10084; FabZ.
 DR TIGRFAMs; TIGR01750; fabZ; 1.
 DR Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
 KW ACT SITE 47 By similarity.
 FT ACT SITE 47
 SQ SEQUENCE 140 AA; 15445 MW; FF12DAED4015226B CRC64;

 Query Match 7.5%; Score 12; DB 1; Length 140;
 Best Local Similarity 100.0%; Pred. No. 0.0016;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 21 LPHRYPMLLVDR 32
 DB 11 LPHRYPMLLVDR 22
 |||||
 RESULT 13
 FABZ_STRPN STANDARD; PRT; 140 AA.
 AC P59201; Q9FBC0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (BC 4.2.1.-)
 DE ((3R)-hydroxymyristoyl ACP dehydratase).
 GN Names=fabZ; OrderedLocNames=SP0424;
 OS *Streptococcus pneumoniae*.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 [1]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP STRAIN=ATCC BAA-334 / TIGR4;
 RC MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M.B., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtzapple E.K., Khouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
 RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*."
 RL Science 293:498-506(2001).
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL; AE007354; AAK74587.1; -; Genomic_DNA.
 DR PIR; B95049; B95049.
 DR HSSP; P18391; IMKA.

DR TIGR; SP0424; -.
 DR HAMAP; MF 00406; -; 1.
 DR InterPro; IPRO10084; FabZ.
 DR TIGRFAMs; TIGR01750; fabZ; 1.
 DR Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
 KW ACT SITE 47 By similarity.
 FT ACT SITE 47
 SQ SEQUENCE 140 AA; 15272 MW; A54650DDA9834813 CRC64;

 Query Match 7.5%; Score 12; DB 1; Length 140;
 Best Local Similarity 100.0%; Pred. No. 0.0016;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 21 LPHRYPMLLVDR 32
 DB 11 LPHRYPMLLVDR 22
 |||||
 RESULT 14
 FABZ_STRR6 STANDARD; PRT; 140 AA.
 AC P59202; Q9FBC0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
 DE ((3R)-hydroxymyristoyl ACP dehydratase).
 GN Names=fabZ; OrderedLocNames=spr0384;
 OS *Streptococcus pneumoniae* (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=171101;
 [1]
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=20365714; PubMed=10910344; DOI=10.1038/35018162;
 RA Heath R.J., Rock C.O.;
 RT "A triclosoan-resistant bacterial enzyme."
 RL Nature 406:145-146(2000).
 [2]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX MEDLINE=21429245; PubMed=11544234;
 RX DOI=10.1128/JB.183.19.5709-5717.2001;
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
 RA DeHoff B.S., Batrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
 RA McAhren S.M., McHenry M., Mcleaster K., Mundy C.W., Nicas T.I.,
 RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
 RA Glass J.I.;
 RT "Genome of the bacterium *Streptococcus pneumoniae* strain R6.";
 RL J. Bacteriol. 183:5709-5717(2001).
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL; AF197933; AAF98278.1; -; Genomic_DNA.
 DR EMBL; AE008418; AAK99188.1; -; Genomic_DNA.
 DR PIR; H97919; H97919.
 DR HSSP; P18391; IMKA.
 DR HAMAP; MF 00406; -; 1.
 DR InterPro; IPRO10084; FabZ.
 DR TIGRFAMs; TIGR01750; fabZ; 1.
 DR Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
 KW ACT SITE 47
 FT ACT SITE 47
 SQ SEQUENCE 133 133
 CONFLICT 133 133
 L -> V (in Ref. 1).
 I -> T (in Ref. 1).

```

SQ SEQUENCE 140 AA; 15298 MW; 6646401713F382B8 CRC64;
Query Match 7.5%; Score 12; DB 1; Length 140;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
Db 11 LPHRYPMLLVDR 22

RESULT 15
FABZ_STRT1
ID FABZ_STRT1 STANDARD; PRT; 141 AA.
AC QSM177;
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=stu0390;
OS Streptococcus thermophilus (strain CNRZ 1066).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
OX NCBI_TaxID=299768;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinkis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goldsman E., Mazur M., Pusch G.D.,
RA Fongstein M., Overbeek R., Kyprides N., Fumelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burtan S., Boutry M., Delcours J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; CP000024; AAV61992.1; -; Genomic_DNA.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 48 48 By similarity.
SQ SEQUENCE 141 AA; 15445 MW; 66E18ED333642A67 CRC64;

Query Match 7.5%; Score 12; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
Db 11 LPHRYPMLLVDR 23

RESULT 16
FABZ_STRT2
ID FABZ_STRT2 STANDARD; PRT; 141 AA.
AC QSM5R4;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=stu0390;
OS Streptococcus thermophilus (strain CNRZ 1066).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
OX NCBI_TaxID=299768;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinkis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goldsman E., Mazur M., Pusch G.D.,
RA Fongstein M., Overbeek R., Kyprides N., Fumelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burtan S., Boutry M., Delcours J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; CP000024; AAV61992.1; -; Genomic_DNA.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 48 48 By similarity.
SQ SEQUENCE 141 AA; 15445 MW; 66E18ED333642A67 CRC64;

Query Match 7.5%; Score 12; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
Db 11 LPHRYPMLLVDR 23

RESULT 17
FABZ_STRT3
ID FABZ_STRT3 STANDARD; PRT; 164 AA.
AC Q8XZ10;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=RS05277;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=GM1000;
RX MEDLINE=11823852; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brothier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002)
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

```

```

OS Streptococcus thermophilus (strain ATCC BAA-250 / LMG 18311).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
OX NCBI_TaxID=264199;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15543133; DOI=10.1038/nbt1034;
RA Bolotin A., Quinkis B., Renault P., Sorokin A., Ehrlich S.D.,
RA Kulakauskas S., Lapidus A., Goldsman E., Mazur M., Pusch G.D.,
RA Fongstein M., Overbeek R., Kyprides N., Fumelle B., Prozzi D.,
RA Ngui K., Masuy D., Hancy F., Burtan S., Boutry M., Delcours J.,
RA Goffeau A., Hols P.;
RT "Complete sequence and comparative genome analysis of the dairy
bacterium Streptococcus thermophilus.";
RL Nat. Biotechnol. 22:1554-1558(2004).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; CP000023; AAV60108.1; -; Genomic_DNA.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 48 48 By similarity.
SQ SEQUENCE 141 AA; 15445 MW; 66E18ED333642A67 CRC64;

Query Match 7.5%; Score 12; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
Db 12 LPHRYPMLLVDR 23

RESULT 17
FABZ_STRT4
ID FABZ_STRT4 STANDARD; PRT; 164 AA.
AC Q8XZ10;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=RS05277;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=GM1000;
RX MEDLINE=11823852; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brothier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002)
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL; AL646064; CAD15117.1; -; Genomic_DNA.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 61 By similarity.
SQ SEQUENCE 164 AA; 18461 MW; 89FAF8127B2A9AE9 CRC64;

Query Match 7.5%; Score 12; DB 1; Length 164;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
DB 24 LPHRYPMLLVDR 35
|||||
-----
RESULT 18
Q7P453_FUSNV PRELIMINARY; PRT; 116 AA.
AC Q7P453;
DT 01-MAR-2004 (T-EMBLrel. 26, Created)
DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
DE (EC 4.2.1.-).
GN Names=FNV0402;
OS Fusobacterium nucleatum subsp. vincentii ATCC 49256.
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=209882;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 49256;
RA Karpatal V., Ivanova N., Anderson I., Reznik G., Bhattacharyya A.,
RA Gardner W.L., Mikhailova N., Larsen N., D'Souza M., Walunas T.,
RA Haselkorn R., Overbeek R., Kyrpides N.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAF01000137; EAA23396.1; -; Genomic_DNA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008610; P:lipid biosynthesis; IEA.
KW Lyase.
SQ SEQUENCE 116 AA; 12801 MW; F50764F28A22FCB3 CRC64;

Query Match 6.9%; Score 11; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 PGVLIVEGMAQ 76
DB 32 PGVLIVEGMAQ 42
|||||
-----
RESULT 19
FABZ_CLOB STANDARD; PRT; 141 AA.
AC Q97DA9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydrase).
GN Names=fabZ; OrderedLocusNames=CAC3571;
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

```

```

OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RD DOI=10.1128/JB.183.16.4823-4838.2001;
RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L.A., Soucaille P.,
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL; AB007854; AAK81494.1; -; Genomic_DNA.
DR PIR; C97338; C97338.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 49 By similarity.
SQ SEQUENCE 141 AA; 15438 MW; 04928B855C2E9F5A CRC64;

Query Match 6.9%; Score 11; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 PHRYPMLLVDR 32
DB 13 PHRYPMLLVDR 23
|||||
-----
RESULT 20
Q4HGA6_CAMCO PRELIMINARY; PRT; 142 AA.
AC Q4HGA6;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Beta-hydroxyacyl-(Acyl-carrier-protein) dehydratase FabZ
DE (EC 4.2.1.-).
GN Names=fabZ; ORFNames=CCO0342;
OS Campylobacter coli RM2228.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306254;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM2228;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R.R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
RT genomes of multiple Campylobacter species.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAF101000003; EAL56875.1; -; Genomic_DNA.
KW Lyase.
SQ SEQUENCE 142 AA; 15937 MW; 2E000F62AC59E479 CRC64;

```



```

Query Match          6.3%; Score 10; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VAAEELKAMI 155
DB 130 VAAEELKAMI 139

RESULT 21
Q4HJP5_CAMLA
ID Q4HJP5_CAMLA PRELIMINARY; PRT; 142 AA.
AC Q4HJP5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Beta-hydroxyacyl-(Acyl-carrier-protein) dehydratase FabZ
( EC 4.2.1.- )
GN Name=fabZ; ORFNames=CLA0287;
OS Campylobacter lari RM2100.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306263;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM2100;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R.R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
RT genomes of multiple Campylobacter species.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFK01000004; EAL54516.1; -; Genomic_DNA.
KW Lyase.
SQ SEQUENCE 142 AA; 16066 MW; D4D1DBA5660BC286 CRC64;

Query Match          6.3%; Score 10; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VAAEELKAMI 155
DB 130 VAAEELKAMI 139

RESULT 22
FABZ_CAMJE
ID FABZ_CAMJE STANDARD; PRT; 146 AA.
AC Q9PIM2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=Cj0273;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204; DOI=10.1038/35001088;
RA Parkhill J., Wren B.W., Mungall K.L., Kelsey J.M., Churcher C.M.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,

Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

DR EMBL; AL139074; CAB72741.1; -; Genomic_DNA.
DR PIR; H81445; H81445.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT SITE 48 By similarity.
SQ SEQUENCE 146 AA; 16440 MW; EPPFA2656B12563A5 CRC64;

Query Match          6.3%; Score 10; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VAAEELKAMI 155
DB 134 VAAEELKAMI 143

RESULT 23
FABZ_CAMJR
ID FABZ_CAMJR STANDARD; PRT; 146 AA.
AC QSHWJ3;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=CJ50322;
OS Campylobacter jejuni (strain RM1221).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=195099;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15660156; DOI=10.1371/journal.pbio.0030015;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Ravel J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural differences and novel potential virulence mechanisms
RT from the genomes of multiple Campylobacter species.";
RL PLOS Biol. 3:72-85(2005).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

DR EMBL; CP000025; AAW34912.1; -; Genomic_DNA.
DR TIGR; CJE0322; -; 1.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.

```

```

FT ACT_SITE 48 48 By similarity.
SQ SEQUENCE 146 AA; 16440 MW; EFA2656B12563A5 CRC64;

Query Match
Best Local Similarity 6.3%; Score 10; DB 1; Length 146;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VAEAEKAMI 155
Db 134 VAEAEKAMI 143

RESULT 24
ID Q4HOA3 CAMUP PRELIMINARY; PRT; 146 AA.
AC Q4HOA3;
DT 13-SEP-2005 (TREMELrel. 31, Created)
DT 13-SEP-2005 (TREMELrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMELrel. 31, Last annotation update)
DE Beta-hydroxyacyl-(Acyl-carrier-protein) dehydratase FabZ
  (EC 4.2.1.-).
GN Name=fabZ; ORFNames=CU1147;
OS Campylobacter upsaliensis RM3195.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306264;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM3195;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R.R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
RT genomes of multiple Campylobacter species.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data
DR EMBL; AAPJ01000007; EAL52872.1; -; Genomic_DNA.
KW Lyase.
SQ SEQUENCE 146 AA; 16591 MW; 8F1CD2BA02891959 CRC64;

Query Match
Best Local Similarity 6.3%; Score 10; DB 2; Length 146;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VAEAEKAMI 155
Db 134 VAEAEKAMI 143

RESULT 25
ID FABZ_THETN STANDARD; PRT; 141 AA.
AC Q8RD71;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydrase).
GN Name=fabZ; Synonyms=fabA; OrderedLocusNames=TTE0175;
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336; DOI=10.1101/gr.219302;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,

```

```

RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of the T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -! FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -! SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -! SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AE012991; AAM23476.1; -; Genomic_DNA.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 47 47 By similarity.
SQ SEQUENCE 141 AA; 15604 MW; 42A5A71CFDF0DBAA CRC64;

Query Match
Best Local Similarity 5.7%; Score 9; DB 1; Length 141;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LLVDRIIEL 36
Db 17 LLVDRIIEL 25

RESULT 26
ID FABZ_PHOLL STANDARD; PRT; 150 AA.
AC Q7N8N6;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydrase).
GN Name=fabZ; OrderedLocusNames=plu0683;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IT01;
RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taouit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose R., Derzelle S., Prevassinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Sigulier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens.";
RL Nat. Biotechnol. 21:1307-1313(2003).
CC -! FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -! SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -! SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BX571861; CAB12978.1; -; Genomic_DNA.
DR PhotolList; plu0683; -; 1.
DR HAMAP; MF_00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 53 53 By similarity.

```

```
SQ SEQUENCE 150 AA; 17079 MW; 6BAF16D204B1F708 CRC64;
Query Match 5.7%; Score 9; DB 1; Length 150;
Best Local Similarity 100.0%; Pred.No.1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFPGLVLI 70
| | | | | | | |
DB 57 KPFPGLVLI 65

RESULT 27
FABZ BORBR
ID FABZ BORBR STANDARD; PRT; 151 AA.
AC Q7WJ83;
DT 05-JUL-2004 (Rel. 44, Created)
DE 05-JUL-2004 (Rel. 44, Last sequence update)
DE 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
DE Name=fabZ; OrderedLocusNames=B2616;
GN Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., Basham D., Bason N., Cherevach I.,
RA Achtman M., Atkin R., Collins M., Cronin A., Davis P., Doggett J.,
RA Chillingworth T., Collins M., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders S., Stevens K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BX640444; CAE33109.1; -; Genomic_DNA.
CC HAMAP; MF_00406; -; 1.
CC InterPro; IPR010084; FabZ.
CC TIGRFAMs; TIGR01750; fabZ; 1.
CC Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 49 49 By similarity.
SQ SEQUENCE 151 AA; 16754 MW; D92AE319C86399E CRC64;

Query Match 5.7%; Score 9; DB 1; Length 151;
Best Local Similarity 100.0%; Pred.No.1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMMLL 29
| | | | | | | |
DB 12 LPHRYPMMLL 20

RESULT 28
FABZ BORBR
ID FABZ BORBR STANDARD; PRT; 151 AA.
AC Q7WJ83;
DT 05-JUL-2004 (Rel. 44, Created)
DE 05-JUL-2004 (Rel. 44, Last sequence update)
DE 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-)
DE ((3R)-hydroxymyristoyl ACP dehydratase).
DE Name=fabZ; OrderedLocusNames=B2616;
GN Bordetella pertussis.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
```

RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Fellwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.";
RL Nat. Genet. 35:32-40(2003).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; BX640415; CAB41720.1; -; Genomic_DNA.
DR HAMAP; MF_00406; -; 1.
DR InterPro; TIGR010084; FapZ.
DR TIGRFAMs; TIGR01750; fapZ; 1.
DR Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
KW ACT SITE 49 49 By similarity.
FT ACT SITE 49 49 By similarity.
SQ SEQUENCE 151 AA; 16754 MW; D992AE319C86399E CRC64;

Query Match 5.7%; Score 9; DB 1; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMILL 29
DB 12 LPHRYPMILL 20
|||||
|

RESULT 30
FAPZ_CANBF STANDARD; PRT; 151 AA.
AC Q7VRD5;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Names=fapZ; OrderedLocusNames=Bfl1282;
OS *Candidatus Blochmannia floridanus*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; ant endosymbionts; *Candidatus Blochmannia*.
OX NCBI_TaxID=203907;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22784745; PubMed=12886019; DOI=10.1073/pnas.1533499100;
RA Gil R., Silva F.J., Zientz E., Delmotte F., Gonzalez-Candelas F.,
RA Latorre A., Rausell C., Kamerbeek J., Gadau J., Hoelldobler B.,
RA van Ham R.C.H.J., Gross R., Moya A.;
RT "The genome sequence of *Blochmannia floridanus*: comparative analysis
RT of reduced genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:9388-9393(2003).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; BX248585; CAD83353.1; -; Genomic_DNA.

DR HAMAP; MF_00406; -; 1.
DR InterPro; TIGR010084; FapZ.
DR TIGRFAMs; TIGR01750; fapZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT SITE 54 54 By similarity.
SQ SEQUENCE 151 AA; 17145 MW; 6C94FF8F89ABED3 CRC64;

Query Match 5.7%; Score 9; DB 1; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIPPGVLLI 70
DB 58 KPIPPGVLLI 66
|||||
|

RESULT 31
FAPZ_ECO57 STANDARD; PRT; 151 AA.
AC POA6Q8; P21774;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Names=fapZ; OrderedLocusNames=z0192; ECs0182;
OS *Escherichia coli* O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; *Escherichia*.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBUNIT: Oligomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- PTM: The N-terminus is blocked (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AB005174; AAGS4482.1; -; Genomic_DNA.
DR EMBL; BA000007; BAB33605.1; -; Genomic_DNA.
DR PIR; F85502; F85502.
DR PIR; F90651; F90651.
DR HAMAP; MF_00406; -; 1.
DR InterPro; TIGR010084; FapZ.
DR TIGRFAMs; TIGR01750; fapZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.

```

FT ACT SITE 54 54 BY similarity.
FT CONFLICT 37 37 F -> L (in Ref. 2).
SQ SEQUENCE 151 AA; 17033 MW; 91F514A7319C2FB1 CRC64;

Query Match 5.7%; Score 9; DB 1; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFPFGVLI 70
Db 58 KPFPFGVLI 66

RESULT 32
FABZ_ECOLI STANDARD; PRT; 151 AA.
AC POA6Q7; P21774;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
DE (3R)-hydroxymyristoyl ACP dehydratase.
GN Name=fabZ; OrderedLocusNames=c0217;
OS Escherichia coli O6
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=O6:HI / CFT073 / ATCC 700928 / UPEC;
RC MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBUNIT: Oligomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- PTM: The N-terminus is blocked (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
DR EMBL; AR016755; AAN78709.1; ALT_INIT; Genomic_DNA.
DR HAMAP; MF 00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRfam; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT SITE 54 54 By similarity.
SQ SEQUENCE 151 AA; 17033 MW; 91F514A7319C2FB1 CRC64;

Query Match 5.7%; Score 9; DB 1; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPFPFGVLI 70
Db 58 KPFPFGVLI 66

RESULT 33
FABZ_ECOLI STANDARD; PRT; 151 AA.
AC POA6Q6; P21774;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
DE (3R)-hydroxymyristoyl ACP dehydratase.
GN Name=fabZ; Synonyms=sefA; OrderedLocusNames=b0180;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=88139188; PubMed=3277952;
RA Coleman J., Raetz C.R.H.;
RT "First committed step of lipid A biosynthesis in Escherichia coli:
RT sequence of the lpxA gene.";
RL J. Bacteriol. 170:1268-1274(1988).
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=93359166; PubMed=8354462; DOI=10.1016/0378-1097(93)90118-L;
RA Postel S.J.;
RT "Purification and characterization of an 'actomyosin' complex from
RT Escherichia coli W3110.";
RL FEMS Microbiol. Lett. 110:295-298(1993).
RN FUNCTION.
RP MEDLINE=95105173; PubMed=7806516;
RA Mohan S., Kelly T.M., Eveland S.S., Raetz C.R.H., Anderson M.S.;
RT "An Escherichia coli gene (fabZ) encoding (3R)-hydroxymyristoyl acyl
RT carrier protein dehydratase. Relation to fabA and suppression of
RT mutations in lipid A biosynthesis.";
RL J. Biol. Chem. 269:32896-32903(1994).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBUNIT: Oligomer.
CC -!- INTERACTION:
CC POA6A8:acpP; NbExp=1; IntAct=EBI-370073; EBI-542566;
CC P36659:cbpA; NbExp=1; IntAct=EBI-370073; EBI-546131;
CC P15716:clpA; NbExp=1; IntAct=EBI-370073; EBI-546140;
CC POA6P9:eno; NbExp=1; IntAct=EBI-370073; EBI-368855;
CC P07019:fffh; NbExp=1; IntAct=EBI-370073; EBI-369938;
CC P00962:glnS; NbExp=1; IntAct=EBI-370073; EBI-550012;
CC P08177:lon; NbExp=1; IntAct=EBI-370073; EBI-547203;
CC P21513:rne; NbExp=1; IntAct=EBI-370073; EBI-549958;

```

```

CC POA7L0:rp1A; NbExp=1; IntAct=EBI-370073, EBI-543771;
CC P60422:rp1B; NbExp=1; IntAct=EBI-370073, EBI-543515;
CC P60438:rp1C; NbExp=1; IntAct=EBI-370073, EBI-542200;
CC P60733:rp1D; NbExp=1; IntAct=EBI-370073, EBI-545597;
CC P63299:rp1E; NbExp=1; IntAct=EBI-370073, EBI-545956;
CC P63390:rp1F; NbExp=1; IntAct=EBI-370073, EBI-546389;
CC POA7R1:rp1I; NbExp=1; IntAct=EBI-370073, EBI-546437;
CC POA7J3:rp1J; NbExp=1; IntAct=EBI-370073, EBI-546827;
CC P02410:rp1M; NbExp=1; IntAct=EBI-370073, EBI-543801;
CC POA7K6:rp1S; NbExp=1; IntAct=EBI-370073, EBI-543891;
CC P61175:rp1V; NbExp=1; IntAct=EBI-370073, EBI-542255;
CC P02349:rp8A; NbExp=1; IntAct=EBI-370073, EBI-546520;
CC POA7V0:rp8B; NbExp=1; IntAct=EBI-370073, EBI-543439;
CC POA7V3:rp8C; NbExp=1; IntAct=EBI-370073, EBI-544851;
CC POA7V8:rp8D; NbExp=1; IntAct=EBI-370073, EBI-543939;
CC POA7M1:rp8E; NbExp=1; IntAct=EBI-370073, EBI-543949;
CC P02359:rp8G; NbExp=1; IntAct=EBI-370073, EBI-543074;
CC P02590:tuFA; NbExp=1; IntAct=EBI-370073, EBI-543077;
CC P77349:ybeZ; NbExp=1; IntAct=EBI-370073, EBI-560255;
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; M19334; AAC36917.1; -; Genomic DNA.
CC EMBL; U00096; AAC73291.1; -; Genomic DNA.
CC EMBL; D83536; BAA7855.1; -; Genomic DNA.
CC EMBL; U70214; AAB08609.1; -; Genomic DNA.
CC PIR; D64742; D64742.
CC IntAct; P0A6Q6; -.
CC EcoBASE; EBI1261; -.
CC EcoGene; EGI1284; fabZ.
CC HAMAP; MF 00406; -; 1.
CC InterPro; IPR010084; FabZ.
CC TIGRFAMS; TIGR01750; fabZ; 1.
CC Complete proteome; Direct protein sequencing; Lipid A biosynthesis;
CC Lipid synthesis; Lyase.
CC ACT SITE 54 54 By similarity.
CC CONFLICT 140 141 EA -> DR (in Ref. 1).
CC CONFLICT 150 151 EA -> RRD (in Ref. 1).
CC SEQUENCE 151 AA; 17033 MW; 91F514A7319C2FB1 CRC64;

Query Match 5.7%; Score 9; DB 1; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIPPGVLI 70
DB 58 KPIPPGVLI 66

RESULT 34
FABZ_ERWCT STANDARD; PRT; 151 AA.
AC Q6D8D2;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-))
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Names=fabZ; OrderedLocustNames=ECAL042;
OS Ewriaia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]-
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RQ STRAIN=SCRI 1043 / ATCC BAA-672;

```

```

RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Heleva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagers K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RA "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BX950851; CAG73953.1; -; Genomic DNA.
CC HAMAP; MF 00406; -; 1.
CC InterPro; IPR010084; FabZ.
CC TIGRFAMS; TIGR01750; fabZ; 1.
CC Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
CC ACT SITE 54 54 By similarity.
CC SEQUENCE 151 AA; 17105 MW; 44A0F1B878069273 CRC64;

Query Match 5.7%; Score 9; DB 1; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIPPGVLI 70
DB 58 KPIPPGVLI 66

RESULT 35
FABZ_SHIFL STANDARD; PRT; 151 AA.
AC P0A6Q9; P21774;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.1.-))
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Names=fabZ; OrderedLocustNames=SF0170, S0173;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]-
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a;
RC MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]-
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RC MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew M.B., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;

```

RT "Complete genome sequence and comparative genomics of Shigella
 RL flexneri serotype 2a strain 2457T";
 CC Infect. Immun. 71:2775-2786(2003).
 CC -|- FUNCTION: Involved in saturated fatty acids biosynthesis.
 CC -|- SUBUNIT: Oligomer (By similarity).
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -|- PTM: The N-terminus is blocked (By similarity).
 CC -|- SIMILARITY: Belongs to the thioester dehydratase family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; AR005674; AAN41832.1; ALT INIT; Genomic_DNA.
 CC EMBL; AR016978; AAP15713.1; -; Genomic_DNA.
 CC HAMAP; MF 00406; -; 1.
 CC InterPro; IPR010084; FabZ.
 CC TIGRFAMs; TIGR01750; fabZ; 1.
 CC Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
 KW ACT SITE 54 By similarity.
 FT ACT SITE 54
 SQ SEQUENCE 151 AA; 17033 MW; 91F514A7319C2PB1 CRC64;
 CC
 CC Query Match 5.7%; Score 9; DB 1; Length 151;
 CC Best Local Similarity 100.0%; Pred. No. 1.8;
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 62 KPFPFGVLI 70
 CC DB 58 KPFPFGVLI 66
 CC
 CC RESULT 36
 CC FABZ WOLSU
 CC ID FABZ WOLSU STANDARD; PRT; 151 AA.
 CC AC Q7MA52;
 CC DT 05-JUL-2004 (Rel. 44, Created)
 CC DT 05-JUL-2004 (Rel. 44, Last sequence update)
 CC DT 10-MAY-2005 (Rel. 47, Last annotation update)
 CC DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
 CC DE ((3R)-hydroxymyristoyl ACP dehydratase).
 CC GN Name=fabZ; OrderedLocusNames=WS0060;
 CC OS Wolinella succinogenes.
 CC OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 CC OC Helicobacteraceae; Wolinella.
 CC OX NCBI_TaxID=844;
 CC [1]
 CC RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 CC RC STRAIN=DSMZ 1740;
 CC RX MEDLINE=22882897; PubMed=14500908; DOI=10.1073/pnas.1932838100;
 CC RA Baar C., Eppinger M., Raddatz G., Simon J., Lanz C., Klimmek O.,
 CC RA Nandakumar R., Gross R., Rosinus A., Keller H., Jagtap P., Linke B.,
 CC RA Meyer F., Lederer H., Schuster S.C.;
 CC RT "Complete genome sequence and analysis of Wolinella succinogenes";
 CC Proc. Natl. Acad. Sci. U.S.A. 100:11690-11695(2003).
 CC -|- FUNCTION: Involved in saturated fatty acids biosynthesis.
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -|- SIMILARITY: Belongs to the thioester dehydratase family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; BX571657; CAE09229.1; -; Genomic_DNA.
 CC HAMAP; MF 00406; -; 1.
 CC InterPro; IPR010084; FabZ.
 CC TIGRFAMs; TIGR01750; fabZ; 1.
 CC Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
 KW ACT SITE 49 By similarity.
 FT ACT SITE 49
 SQ SEQUENCE 151 AA; 17084 MW; 6869BA673438B628 CRC64;

Query Match 5.7%; Score 9; DB 1; Length 151;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 146 VAEAEKAM 154
 DB 137 VAEAEKAM 145
 CC
 CC RESULT 37
 CC FABZ YERP
 CC ID FABZ YERP STANDARD; PRT; 176 AA.
 CC AC Q667K0;
 CC DT 10-MAY-2005 (Rel. 47, Created)
 CC DT 10-MAY-2005 (Rel. 47, Last sequence update)
 CC DT 10-MAY-2005 (Rel. 47, Last annotation update)
 CC DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
 CC DE ((3R)-hydroxymyristoyl ACP dehydratase).
 CC GN Name=fabZ; OrderedLocusNames=YPTB2992;
 CC OS Yersinia pseudotuberculosis.
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC OC Enterobacteriaceae; Yersinia.
 CC OX NCBI_TaxID=633;
 CC [1]
 CC RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 CC RC STRAIN=IP32953 / Serotype I;
 CC RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
 CC RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
 CC RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
 CC RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
 CC RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
 CC RA Derbise A., Hauser L.J., Garcia E.;
 CC RT "Insights into the evolution of Yersinia pestis through whole-genome
 CC comparison with Yersinia pseudotuberculosis";
 CC Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
 CC -|- FUNCTION: Involved in saturated fatty acids biosynthesis.
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -|- SIMILARITY: Belongs to the thioester dehydratase family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; BX936398; CAH22230.1; -; Genomic_DNA.
 CC HAMAP; MF 00406; -; 1.
 CC InterPro; IPR010084; FabZ.
 CC TIGRFAMs; TIGR01750; fabZ; 1.
 CC Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
 KW ACT SITE 54 By similarity.
 FT ACT SITE 54
 SQ SEQUENCE 176 AA; 19631 MW; 3A1602C0A93DD2E2 CRC64;
 CC
 CC Query Match 5.7%; Score 9; DB 1; Length 176;
 CC Best Local Similarity 100.0%; Pred. No. 2;
 CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 62 KPFPFGVLI 70
 CC DB 58 KPFPFGVLI 66
 CC
 CC RESULT 38
 CC FABZ YEREN
 CC ID FABZ YEREN STANDARD; PRT; 178 AA.
 CC AC P32205;
 CC DT 01-OCT-1993 (Rel. 27, Created)
 CC DT 15-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 13-SEP-2005 (Rel. 48, Last annotation update)
 CC DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
 CC DE ((3R)-hydroxymyristoyl ACP dehydratase).
 CC GN Name=fabZ;

OS Yersinia enterocolitica.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=630;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RC STRAIN=EH902;
 RX MEDLINE=94123772; PubMed=8293817; DOI=10.1016/0014-5793(94)80211-4;
 RA Vuorio R., Haerkinen T., Tolvanen M., Vaara M.;
 RT The novel hexapeptide motif found in the acyltransferases LpxA and
 RL LpxD of lipid A biosynthesis is conserved in various bacteria.;
 RT FEBS Lett. 337:289-292(1994).
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 DR EMBL; Z25463; CAA80952.1; ALT_INIT; Genomic_DNA.
 DR PIR; S35968; S35968.
 DR HAMAP; MF_00406; -; 1.
 DR InterPro; IPR010084; FabZ.
 DR TIGRFAMs; TIGR01750; fabZ; 1.
 KW Lipid A biosynthesis; Lipid synthesis; Lyase.
 FT ACT_SITE 54 54 By similarity.
 SQ SEQUENCE 178 AA; 19862 MW; 4C1B6D87847D49BC CRC64;
 Query Match 5.7%; Score 9; DB 1; Length 178;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 62 KPFPFGVLI 70
 Db |||||
 58 KPFPFGVLI 66
 RESULT 39
 FABZ_YERPE STANDARD; PRT; 181 AA.
 ID FABZ_YERPE STANDARD; PRT; 181 AA.
 AC Q8ZH57;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)
 DE ((3R)-hydroxymyristoyl ACP dehydratase).
 GN Name=fabZ; OrderedLocusNames=YPO1055, y3124, YP2795;
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=CO-92 / Biovar Orientalis; DOI=10.1038/35097083;
 RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Feltwell T., Hamlin N., Holroyd S., Jęgleś K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=KIMS / Biovar Mediaevalis;
 RX MEDLINE=21237863; PubMed=12142430;
 RX DOI=10.1128/JB.184.16.4601-4611.2002;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,

RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Nilles R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
 RA Perry R.D.;
 RT "Genome sequence of Yersinia pestis KIM.";
 RL J. Bacteriol. 184:4601-4611(2002).
 [3]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP STRAIN=91001 / Biovar Mediaevalis;
 RX PubMed=15368893;
 RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
 RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Qui B., Qi Z.,
 RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
 RA Yang H., Wang J., Huang P., Yang R.;
 RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
 RT avirulent to humans.";
 RL DNA Res. 11:179-197(2004).
 CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 DR EMBL; AJ414146; CAC89897.1; -; Genomic_DNA.
 DR EMBL; AE013913; AM86674.1; -; Genomic_DNA.
 DR EMBL; AE017137; AAS62979.1; -; Genomic_DNA.
 DR PIR; AF0129; AF0129.
 DR HAMAP; MF_00406; -; 1.
 DR InterPro; IPR010084; FabZ.
 DR TIGRFAMs; TIGR01750; fabZ; 1.
 KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
 FT ACT_SITE 54 54 By similarity.
 FT VARIANT 170 174 Missing (in strain KIMS and strain
 FT 91001).
 SQ SEQUENCE 181 AA; 20170 MW; 584FDE7A62F8C943 CRC64;
 Query Match 5.7%; Score 9; DB 1; Length 181;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 62 KPFPFGVLI 70
 Db |||||
 58 KPFPFGVLI 66
 RESULT 40
 Q8Y0V7_RALSO
 ID Q8Y0V7_RALSO PRELIMINARY; PRT; 298 AA.
 AC Q8Y0V7;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE PUTATIVE PREPHENATE DEHYDROGENASE OXIDOREDUCTASE PROTEIN
 DE (EC 1.3.1.12).
 GN Name=tyra; OrderedLocusNames=RSC0906; ORFNames=RS04509;
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=GMI1000;
 RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Denange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Siguier P., Thebaud P., Whalen M., Wincker P., Levy M.,
 RA Weissbach J., Boucher C.A.;

RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
 RL Nature 415:497-502(2002).
 DR ENBL; AL646061; CAD14608.1; -; Genomic_DNA.
 DR GO; GO:0047794; F:cyclohexadienyl dehydrogenase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0004665; F:prephenate dehydrogenase (NADP+) activity; IEA.
 DR GO; GO:0008977; F:prephenate dehydrogenase activity; IEA.
 DR GO; GO:0006571; P:tyrosine biosynthesis; IEA.
 DR InterPro; IPR008235; PDH fbk insens.
 DR InterPro; IPR003099; Prephen_dehydrog.
 DR Pfam; PF02153; PDH; 1.
 DR PIRSF; PIRSF06786; PDH fbk insens; 1.
 KW Complete proteome; Oxidoreductase.
 SQ SEQUENCE 298 AA; 31444 MW; 8BD3F7C6B00683BD CRC64;

Query Match 5.7%; Score 9; DB 2; Length 298;
 Best Local Similarity 100.0%; Pred. No. 3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 LKAMIAERD 159
 |||||
 Db 261 LKAMIAERD 269

RESULT 41

Q4YGS7 PLABE PRELIMINARY; PRT; 51 AA.
 AC Q4YGS7;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein (Fragment).
 GN ORFNames=PB403260.00.0;
 OS Plasmodium berghei.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5821;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Hall N., Karas M., Raine J.D., Carlton J.M., Kooij T.W.A.,
 RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
 RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
 RA Quall M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
 RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
 RA Jansse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
 RT "A comprehensive survey of the Plasmodium life cycle by genomic,
 RT transcriptomic, and proteomic analyses.";
 RL Science 307:82-86(2005).
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; CAAI01005230; CAI02790.1; -; Genomic_DNA.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 51 AA; 6090 MW; A1187FBD3B43D4A CRC64;

Query Match 5.0%; Score 8; DB 2; Length 51;
 Best Local Similarity 100.0%; Pred. No. 7.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 YHLEVLKH 128
 |||||
 Db 11 YHLEVLKH 18

RESULT 42

Q90X75 GILMI PRELIMINARY; PRT; 94 AA.
 AC Q90X75;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE ABC A transporter (Fragment).
 OS Gillichthys mirabilis (Long-jawed mudsucker).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Gobioidae;
 OC Gobiidae; Gillichthys.
 OK NCBI_TaxID=8222;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Kultz D., Avila K.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF419321; AAL14706.1; -; mRNA.
 FT NON_TER 1
 FT NON_TER 94
 SQ SEQUENCE 94 AA; 11206 MW; A52A5EFA04E984AE CRC64;

Query Match 5.0%; Score 8; DB 2; Length 94;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 IIELOANK 40
 |||||
 Db 64 IIELOANK 71

RESULT 43

Q91404 XENLA PRELIMINARY; PRT; 105 AA.
 AC Q91404;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE XK endo B homolog protein (Fragment).
 GN Name=XK endo B homolog;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA MEDLINE=95290493; PubMed=7772598; DOI=10.1016/0167-4781(95)00049-M;
 RA Adati N., Ito T., Koga C., Kito K., Sakaki Y., Shiohawa K.;
 RT "Differential display analysis of gene expression in developing
 RT embryos of Xenopus laevis.";
 RL Biochim. Biophys. Acta 1262:43-51(1995).
 CC -!- SIMILARITY: Belongs to the intermediate filament family.
 DR EMBL; S78089; AAB34522.1; -; mRNA.
 DR PIR; S55691; S55691.
 DR GO; GO:0005882; C:intermediate filament; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR001664; IF.
 DR Pfam; PF00038; Filament; 1.
 DR PROSITE; PS00226; IF; UNKNOWN_1.
 KW Intermediate filament.
 FT NON_TER 105
 FT NON_TER 105
 SQ SEQUENCE 105 AA; 11896 MW; CDD924315223514A CRC64;

Query Match 5.0%; Score 8; DB 2; Length 105;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 141 VDCKVAAE 148
 |||||
 Db 89 VDCKVAAE 96

RESULT 44

FABZ THEME STANDARD; PRT; 137 AA.
 AC Q9WZQ8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)

```

DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=TW0801;
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;
RX MEDLINE=99287316; PubMed=10360571; DOI=10.1038/20601;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L.A., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.L.,
RA Heidelberg J.F., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RL genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
DR EMBL; AE001747; AAD35883.1; -; Genomic_DNA.
DR FIC; B72335; B72335.
DR TIGR; TW0801; -.
DR HAMAP; MF 00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 46 46 By similarity.
SQ SEQUENCE 137 AA; 15300 MW; ED2FE248F02299D0 CRC64;

Query Match 5.0%; Score 8; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 PGVLIVEG 73
Db 54 PGVLIVEG 61
|||||
|||||

RESULT 45
FABZ SYMTH STANDARD; PRT; 142 AA.
AC O67791;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=STH17;
OS Symbiobacterium thermophilum.
OC Bacteria; Actinobacteria; Symbiobacterium.
OX NCBI_TaxID=2734;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=T / IAM 14863;
RX PubMed=15383646; DOI=10.1093/nar/gkh830;
RA Ueda K., Yamashita A., Ishikawa J., Shimada M., Watsui T.,
RA Morimura K., Ikeda H., Hattori M., Beppu T.;
RT "Genome sequence of Symbiobacterium thermophilum, an uncultivable
RT bacterium that depends on microbial commensalism.";
RL Nucleic Acids Res. 32:4937-4944(2004).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC

```

```

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
DR EMBL; AP006840; BAD39102.1; -; Genomic_DNA.
DR HAMAP; MF 00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 41 41 By similarity.
SQ SEQUENCE 142 AA; 15449 MW; 06C314FD13289981 CRC64;

Query Match 5.0%; Score 8; DB 1; Length 142;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 PIFPGVLI 70
Db 46 PIFPGVLI 53
|||||
|||||

RESULT 46
QSE3F0_VIBF1 PRELIMINARY; PRT; 150 AA.
AC QSE3F0;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (T-EMBLrel. 30, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase
DE (EC 4.2.1.-))
DE OrderedLocusNames=VF1951;
GN Vibrio fischeri (strain ATCC 700601 / ES114).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=312309;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15703294; DOI=10.1073/pnas.0409900102;
RA Ruby E.G., Urbanowski M., Campbell J., Dunn A., Faini M., Gunsalus R.,
RA Lostro P., Lupp C., McCann J., Millikan D., Schaefer A., Stabb E.,
RA Stevens A., Visick K., Whistler C., Greenberg E.P.;
RT "Complete genome sequence of Vibrio fischeri: a symbiotic bacterium
RT with pathogenic congeners.";
RL Proc. Natl. Acad. Sci. U.S.A. 102:3004-3009(2005).
DR EMBL; CP000020; AAW86446.1; -; Genomic_DNA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabZ; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 150 AA; 17033 MW; 9C75F3B016854B1D CRC64;

Query Match 5.0%; Score 8; DB 2; Length 150;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 PIFPGVLI 70
Db 59 PIFPGVLI 66
|||||
|||||

RESULT 47
FABZ_IDILO STANDARD; PRT; 151 AA.
AC QSR022;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabZ; OrderedLocusNames=IL0834;
OS Idiomarina loihiensis.

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Idionarinaceae; Idionarina.
OX NCBI_TaxID=135577;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=12-TR / DSM 15497 / ATCC BAA-735;
RX PubMed=15596722; DOI=10.1073/pnas.0407638102;
RA Hou S., Saw J.H., Lee K.S., Freitas T.A., Belisle C., Kawarabayasi Y.,
RA Donachien S.P., Pikina A., Galperin M.Y., Koonin E.V., Makarova K.S.,
RA Omelchenko M.V., Sorokin A., Wolf Y.I., Li Q.X., Keum Y.S.,
RA Campbell S., Denery J., Aizawa S.-I., Shibata S., Malahoff A.,
RA Alam M.;
RT "Genome sequence of the deep-sea gamma-proteobacterium Idionarina
RT loihensis reveals amino acid fermentation as a source of carbon and
RT energy.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:18036-18041(2004).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AE017340; AAV81674.1; -; Genomic_DNA.
DR HAMAP; MF 00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabz; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 54 54 By similarity.
SQ SEQUENCE 151 AA; 16991 MW; 78464472C98C48A CRC64;

Query Match 5.0%; Score 8; DB 1; Length 151;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 62 KPFFPGVL 69
Db 58 KPFFPGVL 65
|||||

RESULT 48
FABZ PASMU STANDARD; PRT; 152 AA.
AC Q9CJ99;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabz; OrderedLocusNames=PM1995;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; AB006236; AAK04079.1; -; Genomic_DNA.
DR HAMAP; MF 00406; -; 1.
DR InterPro; IPR010084; FabZ.
DR TIGRFAMs; TIGR01750; fabz; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 57 57 By similarity.
SQ SEQUENCE 152 AA; 17291 MW; 7C38DBB9AFC927 CRC64;

Query Match 5.0%; Score 8; DB 1; Length 152;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 63 PIFPGVLI 70
Db 62 PIFPGVLI 69
|||||

RESULT 49
FABZ RHIME STANDARD; PRT; 154 AA.
AC Q92Q46;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ((3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-))
DE ((3R)-hydroxymyristoyl ACP dehydratase).
GN Name=fabz; OrderedLocusNames=R01504; ORFNames=SMC02092;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
RA Capela D., Barloy-Hubier F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -!- FUNCTION: Involved in saturated fatty acids biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the thioester dehydratase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AL591787; CAC46083.1; -; Genomic_DNA.
DR HAMAP; MF 00406; -; 1.
DR InterPro; IPR010084; fabz.
DR TIGRFAMs; TIGR01750; fabz; 1.
KW Complete proteome; Lipid A biosynthesis; Lipid synthesis; Lyase.
FT ACT_SITE 57 57 By similarity.
SQ SEQUENCE 154 AA; 16994 MW; 6E1BB0EA6F8CDF40 CRC64;

Query Match 5.0%; Score 8; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 28 LLVDRIIE 35
Db 27 LLVDRIIE 34
|||||

RESULT 50
Q9TIF3_9CAUD

ID Q9TIF3_9CAUD PRELIMINARY; PRT; 159 AA.
AC Q9TIF3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Lactobacillus bacteriophage phi adh.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=12417;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93231538; PubMed=8472961; DOI=10.1016/0378-1119(93)90590-Y;
RA Frenaux C., De Antoni G., Raya R., Klaenhammer T.;
RT "Genetic organization and sequence of the region encoding integrative
RT functions from Lactobacillus gasseri temperate bacteriophage phi-
RT adh.";
RL Gene 126:61-66(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=95138034; PubMed=7836307;
RA Henrich B., Binshofer B., Blaesi U.;
RT "Primary structure and functional analysis of the lysis genes of
RT Lactobacillus gasseri bacteriophage phi-adh.";
RL J. Bacteriol. 177:723-732(1995).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Engel G., Altermann E., Klein J., Henrich B.;
RT "Structure of a genome region of the Lactobacillus gasseri temperate
RT phage phi adh covering a repressor gene and cognate promoters.";
RL Gene 210:67-70(1998).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99384014; PubMed=10452953; DOI=10.1016/S0378-1119(99)00236-X;
RA Altermann E., Klein J., Henrich B.;
RT "Primary structure and features of the genome of the Lactobacillus
RT gasseri temperate bacteriophage phi-adh.";
RL Gene 236:333-346(1999).
DR EMBL; AJ131519; CAB52524.1; -; Genomic_DNA.
DR InterPro; IPR006967; DUF646_phage.
DR Pfam; PF04883; DUF646; 1.
KW Hypothetical protein.
SQ SEQUENCE 159 AA; 18224 MW; A94886E27E47E50D CRC64;

Query Match 5.0%; Score 8; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VAAEELKA 153
|||
Db 141 VAAEELKA 148

Search completed: January 10, 2006, 13:41:10
Job time : 74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:38:06 ; Search time 23 Seconds
(without alignments)
571.541 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 159

Sequence: 1 MEQSHNLOSQFFIEHILQI.....QVDGKVAEAEUKAMIAERD 159

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 572060 seqs, 82675679 residues

Word size : 8

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/prodata/1/iaa/5 COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/6 COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/H COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/PCUS COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/RE COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	7.5	140	2 US-09-196-388-2	Sequence 2, Appli
2	12	7.5	140	2 US-09-583-110-3906	Sequence 3906, Ap
3	12	7.5	146	2 US-09-107-433-2892	Sequence 2892, Ap
4	12	7.5	160	2 US-10-089-019-18	Sequence 18, Appl
5	9	5.7	162	2 US-09-543-681A-6627	Sequence 6627, Ap

ALIGNMENTS

RESULT 1
US-09-196-388-2

; Sequence 2, Application US/09196388

; Patent No. 6277595

; GENERAL INFORMATION:

; APPLICANT: Warren, Patrick

; APPLICANT: Konstantinidis, Alex

; APPLICANT: Russell, Robert

; TITLE OF INVENTION: fabz

; FILE REFERENCE: GM10182

; CURRENT APPLICATION NUMBER: US/09/196,388

; CURRENT FILING DATE: 1998-11-19

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 140

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae
US-09-196-388-2

Query Match 7.5%; Score 12; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LPHRYPMLLVDR 32
| | | | | | | | | |
Db 11 LPHRYPMLLVDR 22

RESULT 2

US-09-583-110-3906
; Sequence 3906, Application US/09583110

; Patent No. 669703

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al.

; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus

; FILE REFERENCE: PATH00-07A

; CURRENT APPLICATION NUMBER: US/09/583,110

; CURRENT FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/107,433

; PRIOR FILING DATE: 1998-06-30

; PRIOR APPLICATION NUMBER: US 60/085,131

; PRIOR FILING DATE: 1998-05-12

; PRIOR APPLICATION NUMBER: US 60/051,553

; PRIOR FILING DATE: 1997-07-02

; NUMBER OF SEQ ID NOS: 5322

; SEQ ID NO 3906

; LENGTH: 140

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

US-09-583-110-3906

Query Match 7.5%; Score 12; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LPHRYPMLLVDR 32
| | | | | | | | | |
Db 11 LPHRYPMLLVDR 22

RESULT 3

US-09-107-433-2892

; Sequence 2892, Application US/09107433

; Patent No. 6800744

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS
THERAPEUTICS

; NUMBER OF SEQUENCES: 5206

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: <Unknown>

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: <Unknown>

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,433

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/ 085131

; FILING DATE: May 12, 1998

; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 2892:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acids
; LENGTH: 146 amino acids
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...146
; SEQUENCE DESCRIPTION: SEQ ID NO: 2892:
US-09-107-433-2892

Query Match 7.5%; Score 12; DB 2; Length 146;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
Db 17 LPHRYPMLLVDR 28

RESULT 4

US-10-089-019-18
; Sequence 18, Application US/10089019
; Patent No. 6951729
; GENERAL INFORMATION:
; APPLICANT: DEMWOLF, WALTER E. JR
; APPLICANT: KALLENDER, HOWARD
; APPLICANT: LONSDALE, JOHN T.
; TITLE OF INVENTION: METHODS FOR MAKING AND USING FATTY ACID
; TITLE OF INVENTION: SYNTHESIS PATHWAY REAGENTS
; FILE REFERENCE: GM50068
; CURRENT APPLICATION NUMBER: US/10/089,019
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: PCT/US00/29451
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/161,775
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-089-019-18

Query Match 7.5%; Score 12; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
Db 31 LPHRYPMLLVDR 42

RESULT 5

US-09-543-681A-6627
; Sequence 6627, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6627
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6627

Query Match 5.7%; Score 9; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIFPGVLI 70
Db 69 KPIFPGVLI 77

Search completed: January 10, 2006, 13:42:09
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:38:22 ; Search time 61 Seconds

(without alignments)
1089.096 Million cell updates/sec

Title: US-10-662-126-37

Perfect score: 159

Sequence: 1 MEQSHNLSQPFIEHLQI.....QVDGKVAEALKAMIAERD 159

Scoring table: OLIGO /
Gapop 60.0 , Gapext 60.0

Searched: 1867569 seqs, 417829326 residues

Word size : 6

Total number of hits satisfying chosen parameters: 43

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

Database : Published Applications_AA_Main:*
1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	159	100.0	159	4	US-10-282-122A-58982
2	159	100.0	159	5	US-10-662-126-37
3	159	100.0	159	5	US-10-965-006-11
4	118	74.2	159	3	US-09-815-242-11601
5	118	74.2	159	4	US-10-335-977-6975
6	38	23.9	60	4	US-10-335-977-6974
7	12	7.5	139	4	US-10-282-122A-74705
8	12	7.5	140	3	US-09-815-242-13616
9	12	7.5	140	3	US-09-932-702-2
10	12	7.5	140	4	US-10-282-122A-73809
11	12	7.5	140	5	US-10-472-928-692
12	12	7.5	140	5	US-10-965-006-13
13	12	7.5	146	5	US-10-617-320-2892
14	11	6.9	141	4	US-10-282-122A-51842
15	10	6.3	146	4	US-10-282-122A-54269
16	9	5.7	102	4	US-10-282-122A-55351
17	9	5.7	102	4	US-10-282-122A-55353
18	9	5.7	146	5	US-10-650-274-133
19	9	5.7	151	3	US-09-815-243-10045
20	9	5.7	151	4	US-10-369-493-707
21	9	5.7	151	4	US-10-282-122A-42949
22	9	5.7	151	4	US-10-282-122A-51041
23	9	5.7	151	5	US-10-965-006-10
24	9	5.7	157	4	US-10-369-493-207
25	9	5.7	172	4	US-10-369-493-21205
26	9	5.7	172	4	US-10-282-122A-68573
27	9	5.7	181	4	US-10-282-122A-77967

28	8	5.0	133	3	US-09-738-769A-4	Sequence 4, Appli
29	8	5.0	137	4	US-10-369-493-2971	Sequence 2971, Ap
30	8	5.0	139	4	US-10-282-122A-57925	Sequence 57925, A
31	8	5.0	152	4	US-10-282-122A-67466	Sequence 67466, A
32	8	5.0	154	4	US-10-369-493-19450	Sequence 19450, A
33	8	5.0	222	4	US-10-062-254-88	Sequence 88, Appl
34	8	5.0	222	4	US-10-767-701-35176	Sequence 35176, A
35	8	5.0	437	4	US-10-425-114-46464	Sequence 46464, A
36	8	5.0	478	4	US-10-369-493-2615	Sequence 2615, Ap
37	8	5.0	561	4	US-10-437-963-203850	Sequence 203850,
38	8	5.0	562	4	US-10-425-115-231422	Sequence 231422,
39	8	5.0	565	4	US-10-062-254-92	Sequence 92, Appl
40	8	5.0	565	4	US-10-425-115-300838	Sequence 300838,
41	8	5.0	590	4	US-10-437-963-198210	Sequence 198210, A
42	8	5.0	684	4	US-10-156-761-10463	Sequence 10463, A
43	8	5.0	910	4	US-10-369-493-21509	Sequence 21509, A

ALIGNMENTS

RESULT 1

US-10-282-122A-58982
; Sequence 58982, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 58982

; LENGTH: 159

; TYPE: PRT

; ORGANISM: Helicobacter pylori

US-10-282-122A-58982

Query Match 100.0%; Score 159; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 9e-155;

Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEQSHONLQSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEDEVFNHGHP 60
DB 1 MEQSHONLQSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEDEVFNHGHP 60
QY 61 NKPIFPGLVIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
DB 61 NKPIFPGLVIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
QY 121 YHLEVLKHKGMIWQVGTAQVDGKVVAAELKAMIAERD 159
DB 121 YHLEVLKHKGMIWQVGTAQVDGKVVAAELKAMIAERD 159

RESULT 2
US-10-662-126-37
; Sequence 37, Application US/10662126
; Publication No. US20050063987A1
; GENERAL INFORMATION:
; APPLICANT: Knapp, Bernhard
; APPLICANT: Hundt, Erika
; APPLICANT: Schmidt, Karl-Heinz
; TITLE OF INVENTION: Proteins, In Particular Membrane Proteins, of
; FILE REFERENCE: Helicobacter Pylori, Their Preparation and Use
; CURRENT APPLICATION NUMBER: US/10/662,126
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: 09/230,158
; PRIOR FILING DATE: 1999-01-19
; PRIOR APPLICATION NUMBER: PCT/IB97/00981
; PRIOR FILING DATE: 1997-07-25
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; TYPE: PRT
; LENGTH: 159
; ORGANISM: Helicobacter pylori
US-10-662-126-37

Query Match 100.0%; Score 159; DB 5; Length 159;
Best Local Similarity 100.0%; Pred. No. 9e-155;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEQSHONLQSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEDEVFNHGHP 60
DB 1 MEQSHONLQSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEDEVFNHGHP 60
QY 61 NKPIFPGLVIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
DB 61 NKPIFPGLVIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
QY 121 YHLEVLKHKGMIWQVGTAQVDGKVVAAELKAMIAERD 159
DB 121 YHLEVLKHKGMIWQVGTAQVDGKVVAAELKAMIAERD 159

RESULT 3
US-10-965-006-11
; Sequence 11, Application US/10965006
; Publication No. US20050221462A1
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, ALED
; APPLICANT: DHARAMSI, AKIL
; APPLICANT: VEDADI, MASOUD
; APPLICANT: DOMAGALA, MEGAN
; APPLICANT: KIMBER, MATTHEW
; APPLICANT: VALLEE, FRANCOIS
; TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM PSEUDOMONAS AERUGINOSA
; FILE REFERENCE: IPT-257.01
; CURRENT APPLICATION NUMBER: US/10/965,006
; CURRENT FILING DATE: 2004-10-14
; PRIOR APPLICATION NUMBER: PCT/CA03/00560

PRIOR FILING DATE: 2003-04-17
PRIOR APPLICATION NUMBER: 60/373,321
PRIOR FILING DATE: 2002-04-17
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 3.3
SEQ ID NO 11
LENGTH: 159
TYPE: PRT
ORGANISM: Helicobacter pylori
US-10-965-006-11

Query Match 100.0%; Score 159; DB 5; Length 159;
Best Local Similarity 100.0%; Pred. No. 9e-155;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEQSHONLQSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEDEVFNHGHP 60
DB 1 MEQSHONLQSQFFIEHILQILPHRYPMMLVDRIIELOANKKIYVAYKNITFNEDEVFNHGHP 60
QY 61 NKPIFPGLVIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
DB 61 NKPIFPGLVIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVYFMTIDKVKFRIPVTPGDRLE 120
QY 121 YHLEVLKHKGMIWQVGTAQVDGKVVAAELKAMIAERD 159
DB 121 YHLEVLKHKGMIWQVGTAQVDGKVVAAELKAMIAERD 159

RESULT 4
US-09-815-242-11601
; Sequence 11601, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11601
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11601

Query Match 74.2%; Score 118; DB 3; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.1e-112;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 KIVAYKNITFNEDEVFNHGHPNKPFPGLVIVEGMAQTGGFLAFTSLWGFDPPEIAKTKIVY 100

Db 41 KIVAYKNITFNEVDVFNHGFNPKPIPGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTIYV 100
QY 101 FMTIDVKFRIPVTPGDRLEYHLEVLKHKGMIVQVGGTAQVDGKVVAAELKAMIAER 158
Db 101 FMTIDVKFRIPVTPGDRLEYHLEVLKHKGMIVQVGGTAQVDGKVVAAELKAMIAER 158
RESULT 5
US-10-335-977-6975
; Sequence 6975, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 6975:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 159 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...159
; SEQUENCE DESCRIPTION: SEQ ID NO: 6975:
US-10-335-977-6975
Query Match 74.2%; Score 118; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.1e-112;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 41 KIVAYKNITFNEVDVFNHGFNPKPIPGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTIYV 100
Db 41 KIVAYKNITFNEVDVFNHGFNPKPIPGVLIVEGMAQTGGFLAFTSLWGFDPPEIAKTIYV 100
QY 101 FMTIDVKFRIPVTPGDRLEYHLEVLKHKGMIVQVGGTAQVDGKVVAAELKAMIAER 158
Db 101 FMTIDVKFRIPVTPGDRLEYHLEVLKHKGMIVQVGGTAQVDGKVVAAELKAMIAER 158
RESULT 6
US-10-335-977-6974

; Sequence 6974, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 6974:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...60
; SEQUENCE DESCRIPTION: SEQ ID NO: 6974:
US-10-335-977-6974
Query Match 23.9%; Score 38; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 6.3e-31;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 58 HFPNKPFPGLVVEGMAQTGGFLAFTSLWGFDPPEIAK 95
Db 5 HFPNKPFPGLVVEGMAQTGGFLAFTSLWGFDPPEIAK 42
RESULT 7
US-10-282-122A-74705
; Sequence 74705, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.

```
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO: 74705
/ LENGTH: 139
/ TYPE: PRT
/ ORGANISM: Streptococcus pyogenes
US-10-282-122A-74705

Query Match          7.5%; Score 12; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 0.00064;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
DB 10 LPHRYPMLLVDR 21
|||||

RESULT 8
US-09-815-242-13616
/ Sequence 13616, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
```

```
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO: 13616
/ LENGTH: 140
/ TYPE: PRT
/ ORGANISM: Streptococcus pneumoniae
US-09-815-242-13616

Query Match          7.5%; Score 12; DB 3; Length 140;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
DB 11 LPHRYPMLLVDR 22
|||||

RESULT 9
US-09-932-702-2
/ Sequence 2, Application US/09932702
/ Patent No. US20020119520A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Patrick
/ APPLICANT: Konstantinidis, Alex
/ APPLICANT: Russell, Robert
/ TITLE OF INVENTION: fabz
/ FILE REFERENCE: GM10182
/ CURRENT APPLICATION NUMBER: US/09/932,702
/ CURRENT FILING DATE: 2001-08-17
/ PRIOR APPLICATION NUMBER: 09/196,388
/ PRIOR FILING DATE: 1998-11-19
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: FastSEQ for Windows Version 3.0
/ SEQ ID NO: 2
/ LENGTH: 140
/ TYPE: PRT
/ ORGANISM: Streptococcus pneumoniae
US-09-932-702-2

Query Match          7.5%; Score 12; DB 3; Length 140;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
DB 11 LPHRYPMLLVDR 22
|||||

RESULT 10
US-10-282-122A-73809
/ Sequence 73809, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
```

;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/230,335
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/230,347
;; PRIOR FILING DATE: 2000-09-09
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/267,636
;; PRIOR FILING DATE: 2001-02-09
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 78614
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 73809
;; LENGTH: 140
;; TYPE: PRT
;; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73809

Query Match 7.5%; Score 12; DB 4; Length 140;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
Db 11 LPHRYPMLLVDR 22
|||||
|||||

RESULT 11
US-10-472-928-692
;; Sequence 692, Application US/10472928
;; Publication No. US20050020813A1
;; GENERAL INFORMATION:
;; APPLICANT: CHIRON SPA
;; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
;; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
;; FILE REFERENCE: P026926W0
;; CURRENT APPLICATION NUMBER: US/10/472,928
;; CURRENT FILING DATE: 2003-09-26
;; PRIOR APPLICATION NUMBER: GB-0107658.7
;; PRIOR FILING DATE: 2001-03-27
;; NUMBER OF SEQ ID NOS: 4979
;; SOFTWARE: SeqWin99, version 1.03
;; SEQ ID NO 692
;; LENGTH: 140
;; TYPE: PRT
;; ORGANISM: Streptococcus pneumoniae
;; FEATURE:
;; OTHER INFORMATION: (3R)-hydroxymyristoyl-(acyl-carrier-protein) dehydratase (fabZ)
;; OTHER INFORMATION: Cellular location: cytoplasm
;; OTHER INFORMATION: Similar to strain R6 sequence 15902428 (7.E-75)
US-10-472-928-692

Query Match 7.5%; Score 12; DB 5; Length 140;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
Db 11 LPHRYPMLLVDR 22
|||||
|||||

RESULT 12
US-10-965-006-13
;; Sequence 13, Application US/10965006

;; Publication No. US20050221462A1
;; GENERAL INFORMATION:
;; APPLICANT: EDWARDS, ALED
;; APPLICANT: DHARAMSI, AKIL
;; APPLICANT: VEDADI, MASOUD
;; APPLICANT: DOMAGALA, MEGAN
;; APPLICANT: KIMBER, MATTHEW
;; APPLICANT: VALLER, FRANCOIS
;; TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM PSEUDOMONAS AERUGINOSA
;; FILE REFERENCE: IPT-257.01
;; CURRENT APPLICATION NUMBER: US/10/965,006
;; CURRENT FILING DATE: 2004-10-14
;; PRIOR APPLICATION NUMBER: PCT/CA03/00560
;; PRIOR FILING DATE: 2003-04-17
;; PRIOR APPLICATION NUMBER: 60/373,321
;; PRIOR FILING DATE: 2002-04-17
;; NUMBER OF SEQ ID NOS: 15
;; SOFTWARE: PatentIn Ver. 3.3
;; SEQ ID NO 13
;; LENGTH: 140
;; TYPE: PRT
;; ORGANISM: Streptococcus pneumoniae
US-10-965-006-13

Query Match 7.5%; Score 12; DB 5; Length 140;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPMLLVDR 32
Db 11 LPHRYPMLLVDR 22
|||||
|||||

RESULT 13
US-10-617-320-2892
;; Sequence 2892, Application US/10617320
;; Publication No. US20050136404A1
;; GENERAL INFORMATION:
;; APPLICANT: Lynn A Doucette-Stamm and David Bush
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
;; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
;; THERAPEUTICS
;; NUMBER OF SEQUENCES: 5206
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;; STREET: 100 Beaver Street
;; CITY: Waltham
;; STATE: Massachusetts
;; COUNTRY: USA
;; ZIP: 02354
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: CD-ROM ISO9660
;; COMPUTER: <Unknown>
;; OPERATING SYSTEM: <Unknown>
;; SOFTWARE: <Unknown>
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/617,320
;; FILING DATE: 10-Jul-2003
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/09/107,433
;; FILING DATE: 30-Jun-1998
;; APPLICATION NUMBER: 60/ 085131
;; FILING DATE: May 12, 1998
;; APPLICATION NUMBER: 60/051553
;; FILING DATE: July 2, 1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ariniello, Pamela Deneke
;; REGISTRATION NUMBER: 40,489
;; REFERENCE/DOCKET NUMBER: GTC-011
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (781)893-5007
;; TELEFAX: (781)893-8277
;; INFORMATION FOR SEQ ID NO: 2892:

SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...146
SEQUENCE DESCRIPTION: SEQ ID NO: 2892:
US-10-617-320-2892

Query Match 7.5%; Score 12; DB 5; Length 146;
Best Local Similarity 100.0%; Pred. No. 0.00067; Mismatches 0; Indels 0; Gaps 0;
Matches 12; Conservative 0;

QY 21 LPHRYPMLLVDR 32
|||||
DB 17 LPHRYPMLLVDR 28

RESULT 14

US-10-282-122A-51842
; Sequence 51842, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 51842
LENGTH: 141
TYPE: PRT
ORGANISM: Clostridium acetobutylicum

US-10-282-122A-51842

Query Match 6.9%; Score 11; DB 4; Length 141;
Best Local Similarity 100.0%; Pred. No. 0.0069; Mismatches 11; Conservative 0; Indels 0; Gaps 0;

QY 22 PHRYPMLLVDR 32
|||||
DB 13 PHRYPMLLVDR 23

RESULT 15

US-10-282-122A-54269
; Sequence 54269, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 54269
LENGTH: 146
TYPE: PRT
ORGANISM: Campylobacter jejuni

US-10-282-122A-54269

Query Match 6.3%; Score 10; DB 4; Length 146;
Best Local Similarity 100.0%; Pred. No. 0.076; Mismatches 10; Conservative 0; Indels 0; Gaps 0;

QY 146 VAEAEKXAMI 155
|||||
DB 134 VAEAEKXAMI 143

RESULT 16

US-10-282-122A-55351
; Sequence 55351, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 55351
LENGTH: 102
TYPE: PRT
ORGANISM: Enterobacter cloacae
US-10-282-122A-55351

Query Match 5.7%; Score 9; DB 4; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KPFPFGVLI 70
Db 74 KPFPFGVLI 82

RESULT 17
US-10-282-122A-55353
Sequence 55353, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 55353
LENGTH: 102
TYPE: PRT
ORGANISM: Enterobacter cloacae
US-10-282-122A-55353

Query Match 5.7%; Score 9; DB 4; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KPFPFGVLI 70
Db 74 KPFPFGVLI 82

RESULT 18
US-10-650-274-153
Sequence 153, Application US/10650274
Publication No. US20050202437A1
GENERAL INFORMATION:
APPLICANT: GLENN, MATTHEW
APPLICANT: HAVUKKALA, ILKKA J
APPLICANT: LUBBERS, MARK WILLIAM
APPLICANT: DEKKER, JAMES
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES,
TITLE OF INVENTION: MATERIALS INCORPORATING THEM, AND METHODS FOR USING
TITLE OF INVENTION: THEM.
FILE REFERENCE: 11000.1073
CURRENT APPLICATION NUMBER: US/10/650,274
CURRENT FILING DATE: 2003-08-28
NUMBER OF SEQ ID NOS: 187
SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0
SEQ ID NO 153
LENGTH: 146
TYPE: PRT
ORGANISM: LACTOBACILLUS RHAMNOSUS
US-10-650-274-153

Query Match 5.7%; Score 9; DB 5; Length 146;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 ILPHRYPML 28
Db 16 ILPHRYPML 24

RESULT 19

```
US-09-815-242-10045
; Sequence 10045, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10045
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10045

Query Match          5.7%; Score 9; DB 3; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches          9; Conservative          0; Mismatches          0; Indels          0; Gaps          0;

QY          62 KPIFPGLI 70
Db          58 KPIFPGLI 66

RESULT 20
US-10-369-493-707
; Sequence 707, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 707
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-707

Query Match          5.7%; Score 9; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches          9; Conservative          0; Mismatches          0; Indels          0; Gaps          0;

QY          62 KPIFPGLI 70
Db          58 KPIFPGLI 66

US-09-815-242-10045
; Sequence 10045, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10045
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10045

Query Match          5.7%; Score 9; DB 3; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches          9; Conservative          0; Mismatches          0; Indels          0; Gaps          0;

QY          62 KPIFPGLI 70
Db          58 KPIFPGLI 66

RESULT 20
US-10-369-493-707
; Sequence 707, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 707
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-369-493-707

Query Match          5.7%; Score 9; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches          9; Conservative          0; Mismatches          0; Indels          0; Gaps          0;

QY          62 KPIFPGLI 70
Db          58 KPIFPGLI 66

US-10-282-122A-42949
; Sequence 42949, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42949
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-42949

Query Match          5.7%; Score 9; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches          9; Conservative          0; Mismatches          0; Indels          0; Gaps          0;

QY          62 KPIFPGLI 70
Db          58 KPIFPGLI 66

RESULT 22
US-10-282-122A-51041
; Sequence 51041, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
```

APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
SOFTWARE: PatentIn version 3.1
SEQ ID NO 51041
LENGTH: 151
TYPE: PRT
ORGANISM: Bordetella pertussis
US-10-282-122A-51041

Query Match 5.7%; Score 9; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LPHRYPMLL 29
Db 12 LPHRYPMLL 20

RESULT 23
US-10-965-006-10
Sequence 10, Application US/10965006
Publication No. US20050221462A1
GENERAL INFORMATION:
APPLICANT: EDWARDS, ALED
APPLICANT: DHARAMSI, AKIL
APPLICANT: VEDADI, MASOUD
APPLICANT: DOMAGALA, MEGAN
APPLICANT: KIMBER, MATTHEW
APPLICANT: VALLÉE, FRANÇOIS
TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM PSEUDOMONAS AERUGINOSA
FILE REFERENCE: IPT-257.01
CURRENT APPLICATION NUMBER: US/10/965,006
CURRENT FILING DATE: 2004-10-14
PRIOR APPLICATION NUMBER: PCT/CA03/00560
PRIOR FILING DATE: 2003-04-17
PRIOR APPLICATION NUMBER: 60/373,321
PRIOR FILING DATE: 2002-04-17
NUMBER OF SEQ ID NOS: 15

SOFTWARE: PatentIn Ver. 3.3
SEQ ID NO 10
LENGTH: 151
TYPE: PRT
ORGANISM: Escherichia coli
US-10-965-006-10

Query Match 5.7%; Score 9; DB 5; Length 151;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KPIPPGVLI 70
Db 58 KPIPPGVLI 66

RESULT 24
US-10-369-493-207
Sequence 207, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 207
LENGTH: 157
TYPE: PRT
ORGANISM: Xenorhabdus nematophilus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(157)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-207

Query Match 5.7%; Score 9; DB 4; Length 157;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 KPIPPGVLI 70
Db 64 KPIPPGVLI 72

RESULT 25
US-10-369-493-21205
Sequence 21205, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 21205
LENGTH: 172

; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(172)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-21205

Query Match 5.7%; Score 9; DB 4; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.93;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIFPGLI 70
Db 79 KPIFPGLI 87

RESULT 26
US-10-282-122A-68573
; Sequence 68573, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68573

; LENGTH: 172
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (3)..(3)
; OTHER INFORMATION: X-any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (12)..(12)
; OTHER INFORMATION: X-any amino acid

US-10-282-122A-68573

Query Match 5.7%; Score 9; DB 4; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.93;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIFPGLI 70
Db 81 KPIFPGLI 89

RESULT 27
US-10-282-122A-77967
; Sequence 77967, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77967

; LENGTH: 181
; TYPE: PRT
; ORGANISM: Yersinia pestis
US-10-282-122A-77967

Query Match 5.7%; Score 9; DB 4; Length 181;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 KPIFPGLI 70
Db 58 KPIFPGLI 66

RESULT 28
US-09-738-769A-4
; Sequence 4, Application US/09738769A


```
; Patent No. US20020055148A1
; GENERAL INFORMATION:
; APPLICANT: HONG, Kyongsu
; TITLE OF INVENTION: A CLONING METHOD FOR DNA FRAGMENTS USING ARBITRARILY PRIMED PCR
; FILE REFERENCE: P20153
; CURRENT APPLICATION NUMBER: US/09/738,769A
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 064432/99
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: 09/519,581
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Streptococcus zooepidemicus
US-09-738-769A-4

Query Match          5.0%; Score 8; DB 3; Length 133;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      26 PMLLVDR1 33
      |||||
Db      9 PMLLVDR1 16

RESULT 29
US-10-369-493-2971
; Sequence 2971, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2971
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Thermotoga maritima
US-10-369-493-2971

Query Match          5.0%; Score 8; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      66 PGVLIVEG 73
      |||||
Db      54 PGVLIVEG 61

RESULT 30
US-10-282-122A-57925
; Sequence 57925, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
```

```
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57925
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Enterococcus faecium
US-10-282-122A-57925

Query Match          5.0%; Score 8; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      22 PHRYPMLL 29
      |||||
Db      11 PHRYPMLL 18

RESULT 31
US-10-282-122A-67466
; Sequence 67466, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
```

; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 67466
 ; LENGTH: 152
 ; TYPE: PRT
 ; ORGANISM: Pasteurella multocida
 US-10-282-122A-67466

Query Match 5.0%; Score 8; DB 4; Length 152;
 Best Local Similarity 100.0%; Pred. No. 8.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 PIPPGVLI 70
 |||||
 Db 62 PIPPGVLI 69

RESULT 32
 US-10-369-493-19450
 ; Sequence 19450, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(S2052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 19450
 ; LENGTH: 154
 ; TYPE: PRT
 ; ORGANISM: Myxococcus xanthus
 US-10-369-493-19450

Query Match 5.0%; Score 8; DB 4; Length 154;
 Best Local Similarity 100.0%; Pred. No. 9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LPHRYPML 28
 |||||
 Db 17 LPHRYPML 24

RESULT 33
 US-10-062-254-88
 ; Sequence 88, Application US/10062254
 ; Publication No. US2002013882A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Caboon, Edgar B
 ; APPLICANT: Cahoon, Rebecca E

; APPLICANT: Falco, Saverio Carl
 ; APPLICANT: Fang, Yiwen
 ; APPLICANT: Hantke, Sabine S.
 ; APPLICANT: Lee, Jian-Ming
 ; APPLICANT: Li, Zhongsen
 ; APPLICANT: Miao, Gao-Hua
 ; APPLICANT: Morgante, Michele
 ; APPLICANT: Niu, Xiping
 ; APPLICANT: Odell, Joan
 ; APPLICANT: Rafalski, Antoni
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Zheng, Peizhong
 ; APPLICANT: Zhu, Qun

; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
 ; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/062,254
 ; CURRENT FILING DATE: 2002-02-01
 ; PRIOR APPLICATION NUMBER: 09/630,346
 ; PRIOR FILING DATE: 2000-07-28
 ; PRIOR APPLICATION NUMBER: 60/146511
 ; PRIOR FILING DATE: 1999-07-30
 ; PRIOR APPLICATION NUMBER: 60/156006
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: 60/156899
 ; PRIOR FILING DATE: 1999-09-30
 ; PRIOR APPLICATION NUMBER: 60/157287
 ; PRIOR FILING DATE: 1999-10-01
 ; PRIOR APPLICATION NUMBER: 60/169767
 ; PRIOR FILING DATE: 1999-12-09
 ; PRIOR APPLICATION NUMBER: 60/171054
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: 60/172958
 ; PRIOR FILING DATE: 1999-12-21
 ; PRIOR APPLICATION NUMBER: 60/171515
 ; PRIOR FILING DATE: 1999-12-22
 ; PRIOR APPLICATION NUMBER: 60/173535
 ; PRIOR FILING DATE: 1999-12-29
 ; NUMBER OF SEQ ID NOS: 375
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 88
 ; LENGTH: 222
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (4)
 ; NAME/KEY: UNSURE
 ; LOCATION: (9)
 ; NAME/KEY: UNSURE
 ; LOCATION: (18)
 ; NAME/KEY: UNSURE
 ; LOCATION: (26)
 ; NAME/KEY: UNSURE
 ; LOCATION: (29)
 ; NAME/KEY: UNSURE
 ; LOCATION: (52)
 US-10-062-254-88

Query Match 5.0%; Score 8; DB 4; Length 222;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GFDPEIAK 95
 |||||
 Db 156 GFDPEIAK 163

RESULT 34
 US-10-767-701-35176
 ; Sequence 35176, Application US/10767701
 ; Publication No. US20040172684A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 35176
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C60519_1.pep
US-10-767-701-35176

Query Match          5.0%; Score 8; DB 4; Length 222;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GFDPEIAK 95
Db 95 GFDPEIAK 102

RESULT 35
US-10-425-114-46464
; Sequence 46464, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46464
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700209195_FLI.pep
US-10-425-114-46464

Query Match          5.0%; Score 8; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GFDPEIAK 95
Db 160 GFDPEIAK 167

RESULT 36
US-10-369-493-2615
; Sequence 2615, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
```

```
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2615
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Synechocystis sp.
US-10-369-493-2615
```

```
Query Match          5.0%; Score 8; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GFDPEIAK 95
Db 219 GFDPEIAK 236
```

```
RESULT 37
US-10-437-963-203850
; Sequence 203850, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 203850
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_98995C.1.pep
US-10-437-963-203850
```

```
Query Match          5.0%; Score 8; DB 4; Length 561;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 88 GFDPEIAK 95
Db 284 GFDPEIAK 291
```

```
RESULT 38
US-10-425-115-231422
; Sequence 231422, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 231422
; LENGTH: 562
; TYPE: PRT
```

```
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_142653C.1.pep
US-10-425-115-231422

Query Match          5.0%; Score 8; DB 4; Length 562;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      88 GFDPEIAK 95
Db      285 GFDPEIAK 292

RESULT 39
US-10-062-254-92
; Sequence 92, Application US/10062254
; Publication No. US20020138882A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B
; APPLICANT: Cahoon, Rebecca E
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Fang, Yiwen
; APPLICANT: Hantke, Sabine S.
; APPLICANT: Lee, Jian-Ming
; APPLICANT: Li, Zhongsen
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Morgante, Michele
; APPLICANT: Niu, Xiping
; APPLICANT: Odell, Joan
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; APPLICANT: Zheng, Peizhong
; APPLICANT: Zhu, Qun
; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/062,254
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/630,346
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/146511
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/156006
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/156899
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 60/157287
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/169767
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/171054
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/172958
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/171515
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 60/173535
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 375
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 92
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Zea mays
US-10-062-254-92

Query Match          5.0%; Score 8; DB 4; Length 565;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      88 GFDPEIAK 95
Db      288 GFDPEIAK 295

us-10-662-126-37.oli8.rapbm
```

```
RESULT 40
US-10-425-115-300838
; Sequence 300838, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 300838
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_37433C.1.pep
US-10-425-115-300838

Query Match          5.0%; Score 8; DB 4; Length 565;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      88 GFDPEIAK 95
Db      288 GFDPEIAK 295
```

```
RESULT 41
US-10-437-963-198210
; Sequence 198210, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Bing
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 198210
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_93893C.1.pep
US-10-437-963-198210
```

```
Query Match          5.0%; Score 8; DB 4; Length 590;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      88 GFDPEIAK 95
Db      290 GFDPEIAK 297
```

```
RESULT 42
US-10-156-761-10463
; Sequence 10463, Application US/10156761
```

```
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10463
; LENGTH: 684
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10463

Query Match      5.0%; Score 8; DB 4; Length 684;
Best Local Similarity 100.0%; Pred. No. 34;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      78 GGFLAFTS 85
DB      293 GGFLAFTS 300

RESULT 43
US-10-369-493-21509
; Sequence 21509, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21509
; LENGTH: 910
; TYPE: PRT
; ORGANISM: Methanobacterium thermoautotrophicum
US-10-369-493-21509

Query Match      5.0%; Score 8; DB 4; Length 910;
Best Local Similarity 100.0%; Pred. No. 43;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      147 AEAEELKAM 154
DB      424 AEAEELKAM 431

Search completed: January 10, 2006, 13:43:21
Job time : 61 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2006, 13:39:52 ; Search time 9 Seconds
(without alignments)
150.054 Million cell updates/sec

Title: US-10-662-126-37
Perfect score: 159
Sequence: 1 MEQSHONLQSPFIEHLQI.....QVDGKVVAEAEKAMIAERD 159

Scoring table: OLIGO
Gapop: 60.0 , Gapext 60.0

Searched: 61141 seqs, 8493638 residues

Word size : 8
Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 500 summaries

- Database : Published Applications_AA_New.*
- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description

No matches found

Search completed: January 10, 2006, 13:43:40
Job time : 9 secs